Annals of Mathematics, **142** (1995), 443–551

Modular elliptic curves and Fermat's Last Theorem

By ANDREW WILES*

For Nada, Clare, Kate and Olivia

Cubum autem in duos cubos, aut quadratoquadratum in duos quadratoquadratos, et generaliter nullam in infinitum ultra quadratum potestatem in duos ejusdem nominis fas est dividere: cujus rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

Pierre de Fermat

Introduction

An elliptic curve over **Q** is said to be modular if it has a finite covering by a modular curve of the form $X_0(N)$. Any such elliptic curve has the property that its Hasse-Weil zeta function has an analytic continuation and satisfies a functional equation of the standard type. If an elliptic curve over **Q** with a given j-invariant is modular then it is easy to see that all elliptic curves with the same j-invariant are modular (in which case we say that the j-invariant is modular). A well-known conjecture which grew out of the work of Shimura and Taniyama in the 1950's and 1960's asserts that every elliptic curve over **Q** is modular. However, it only became widely known through its publication in a paper of Weil in 1967 [We] (as an exercise for the interested reader!), in which, moreover, Weil gave conceptual evidence for the conjecture. Although it had been numerically verified in many cases, prior to the results described in this paper it had only been known that finitely many j-invariants were modular.

In 1985 Frey made the remarkable observation that this conjecture should imply Fermat's Last Theorem. The precise mechanism relating the two was formulated by Serre as the ε-conjecture and this was then proved by Ribet in the summer of 1986. Ribet's result only requires one to prove the conjecture for semistable elliptic curves in order to deduce Fermat's Last Theorem.

*The work on this paper was supported by an NSF grant.

Our approach to the study of elliptic curves is via their associated Galois representations. Suppose that ρ_p is the representation of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ on the p-division points of an elliptic curve over $\mathbf{Q}$, and suppose for the moment that ρ_3 is irreducible. The choice of 3 is critical because a crucial theorem of Langlands and Tunnell shows that if ρ_3 is irreducible then it is also modular. We then proceed by showing that under the hypothesis that ρ_3 is semistable at 3, together with some milder restrictions on the ramification of ρ_3 at the other primes, every suitable lifting of ρ_3 is modular. To do this we link the problem, via some novel arguments from commutative algebra, to a class number problem of a well-known type. This we then solve with the help of the paper [TW]. This suffices to prove the modularity of E as it is known that E is modular if and only if the associated 3-adic representation is modular.

The key development in the proof is a new and surprising link between two strong but distinct traditions in number theory, the relationship between Galois representations and modular forms on the one hand and the interpretation of special values of L-functions on the other. The former tradition is of course more recent. Following the original results of Eichler and Shimura in the 1950's and 1960's the other main theorems were proved by Deligne, Serre and Langlands in the period up to 1980. This included the construction of Galois representations associated to modular forms, the refinements of Langlands and Deligne (later completed by Carayol), and the crucial application by Langlands of base change methods to give converse results in weight one. However with the exception of the rather special weight one case, including the extension by Tunnell of Langlands' original theorem, there was no progress in the direction of associating modular forms to Galois representations. From the mid 1980's the main impetus to the field was given by the conjectures of Serre which elaborated on the ε-conjecture alluded to before. Besides the work of Ribet and others on this problem we draw on some of the more specialized developments of the 1980's, notably those of Hida and Mazur.

The second tradition goes back to the famous analytic class number formula of Dirichlet, but owes its modern revival to the conjecture of Birch and Swinnerton-Dyer. In practice however, it is the ideas of Iwasawa in this field on which we attempt to draw, and which to a large extent we have to replace. The principles of Galois cohomology, and in particular the fundamental theorems of Poitou and Tate, also play an important role here.

The restriction that ρ_3 be irreducible at 3 is bypassed by means of an intriguing argument with families of elliptic curves which share a common ρ_5. Using this, we complete the proof that all semistable elliptic curves are modular. In particular, this finally yields a proof of Fermat's Last Theorem. In addition, this method seems well suited to establishing that all elliptic curves over $\mathbf{Q}$ are modular and to generalization to other totally real number fields.

Now we present our methods and results in more detail.

Let f be an eigenform associated to the congruence subgroup $\Gamma_1(N)$ of $\mathrm{SL}_2(\mathbf{Z})$ of weight $k \geq 2$ and character χ. Thus if T_n is the Hecke operator associated to an integer n there is an algebraic integer $c(n, f)$ such that $T_n f = c(n, f) f$ for each n. We let K_f be the number field generated over $\mathbf{Q}$ by the $\{c(n, f)\}$ together with the values of χ and let $\mathcal{O}_f$ be its ring of integers. For any prime λ of $\mathcal{O}_f$ let $\mathcal{O}_{f,\lambda}$ be the completion of $\mathcal{O}_f$ at λ. The following theorem is due to Eichler and Shimura (for $k = 2$) and Deligne (for $k > 2$). The analogous result when $k = 1$ is a celebrated theorem of Serre and Deligne but is more naturally stated in terms of complex representations. The image in that case is finite and a converse is known in many cases.

THEOREM 0.1. *For each prime $p \in \mathbf{Z}$ and each prime $\lambda \mid p$ of $\mathcal{O}_f$ there is a continuous representation*

$$\rho_{f,\lambda}\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathcal{O}_{f,\lambda})$$

which is unramified outside the primes dividing Np and such that for all primes $q \nmid Np$,

$$\operatorname{trace} \rho_{f,\lambda}(\operatorname{Frob} q) = c(q, f), \qquad \det \rho_{f,\lambda}(\operatorname{Frob} q) = \chi(q) q^{k-1}.$$

We will be concerned with trying to prove results in the opposite direction, that is to say, with establishing criteria under which a λ-adic representation arises in this way from a modular form. We have not found any advantage in assuming that the representation is part of a compatible system of λ-adic representations except that the proof may be easier for some λ than for others.

Assume

$$\rho_0 : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\bar{\mathbf{F}}_p)$$

is a continuous representation with values in the algebraic closure of a finite field of characteristic p and that $\det \rho_0$ is odd. We say that ρ_0 is modular if ρ_0 and $\rho_{f,\lambda} \bmod \lambda$ are isomorphic over $\bar{\mathbf{F}}_p$ for some f and λ and some embedding of $\mathcal{O}_f/\lambda$ in $\bar{\mathbf{F}}_p$. Serre has conjectured that every irreducible ρ_0 of odd determinant is modular. Very little is known about this conjecture except when the image of ρ_0 in $\mathrm{PGL}_2(\bar{\mathbf{F}}_p)$ is dihedral, A_4 or S_4. In the dihedral case it is true and due (essentially) to Hecke, and in the A_4 and S_4 cases it is again true and due primarily to Langlands, with one important case due to Tunnell (see Theorem 5.1 for a statement). More precisely these theorems actually associate a form of weight one to the corresponding complex representation but the versions we need are straightforward deductions from the complex case. Even in the reducible case not much is known about the problem in the form we have described it, and in that case it should be observed that one must also choose the lattice carefully as only the semisimplification of $\overline{\rho_{f,\lambda}} = \rho_{f,\lambda} \bmod \lambda$ is independent of the choice of lattice in $K_{f,\lambda}^2$.

If $\mathcal{O}$ is the ring of integers of a local field (containing $\mathbf{Q}_p$) we will say that $\rho : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathcal{O})$ is a lifting of ρ_0 if, for a specified embedding of the residue field of $\mathcal{O}$ in $\bar{\mathbf{F}}_p$, $\bar{\rho}$ and ρ_0 are isomorphic over $\bar{\mathbf{F}}_p$. Our point of view will be to assume that ρ_0 is modular and then to attempt to give conditions under which a representation ρ lifting ρ_0 comes from a modular form in the sense that $\rho \simeq \rho_{f,\lambda}$ over $\overline{K_{f,\lambda}}$ for some f, λ. We will restrict our attention to two cases:

(I) ρ_0 is ordinary (at p) by which we mean that there is a one-dimensional subspace of $\bar{\mathbf{F}}_p^2$, stable under a decomposition group at p and such that the action on the quotient space is unramified and distinct from the action on the subspace.

(II) ρ_0 is flat (at p), meaning that as a representation of a decomposition group at p, ρ_0 is equivalent to one that arises from a finite flat group scheme over $\mathbf{Z}_p$, and $\det \rho_0$ restricted to an inertia group at p is the cyclotomic character.

We say similarly that ρ is ordinary (at p) if, viewed as a representation to $\bar{\mathbf{Q}}_p^2$, there is a one-dimensional subspace of $\bar{\mathbf{Q}}_p^2$ stable under a decomposition group at p and such that the action on the quotient space is unramified.

Let $\varepsilon : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathbf{Z}_p^\times$ denote the cyclotomic character. Conjectural converses to Theorem 0.1 have been part of the folklore for many years but have hitherto lacked any evidence. The critical idea that one might dispense with compatible systems was already observed by Drinfeld in the function field case [Dr]. The idea that one only needs to make a geometric condition on the restriction to the decomposition group at p was first suggested by Fontaine and Mazur. The following version is a natural extension of Serre's conjecture which is convenient for stating our results and is, in a slightly modified form, the one proposed by Fontaine and Mazur. (In the form stated this incorporates Serre's conjecture. We could instead have made the hypothesis that ρ_0 is modular.)

CONJECTURE. *Suppose that* $\rho\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathcal{O})$ *is an irreducible lifting of* ρ_0 *and that* ρ *is unramified outside of a finite set of primes. There are two cases:*

(i) *Assume that* ρ_0 *is ordinary. Then if* ρ *is ordinary and* $\det \rho = \varepsilon^{k-1}\chi$ *for some integer* $k \geq 2$ *and some* χ *of finite order,* ρ *comes from a modular form.*

(ii) *Assume that* ρ_0 *is flat and that* p *is odd. Then if* ρ *restricted to a decomposition group at* p *is equivalent to a representation on a* p*-divisible group, again* ρ *comes from a modular form.*

In case (ii) it is not hard to see that if the form exists it has to be of weight 2; in (i) of course it would have weight k. One can of course enlarge this conjecture in several ways, by weakening the conditions in (i) and (ii), by considering other number fields in place of $\mathbf{Q}$ and by considering groups other than GL_2.

We prove two results concerning this conjecture. The first includes the hypothesis that ρ_0 is modular. Here and for the rest of the paper we will assume that p is an odd prime.

THEOREM 0.2. *Suppose that ρ_0 is irreducible and satisfies either* (I) *or* (II) *above. Suppose also that ρ_0 is modular and that*

(i) *ρ_0 is absolutely irreducible when restricted to $\mathbf{Q}\left(\sqrt{(-1)^{\frac{p-1}{2}}p}\right)$.*

(ii) *If $q \equiv -1 \bmod p$ is ramified in ρ_0 then either $\rho_0|_{D_q}$ is reducible over the algebraic closure where D_q is a decomposition group at q or $\rho_0|_{I_q}$ is absolutely irreducible where I_q is an inertia group at q.*

Then any representation ρ as in the conjecture does indeed come from a modular form.

The only condition which really seems essential to our method is the requirement that ρ_0 be modular.

The most interesting case at the moment is when $p = 3$ and ρ_0 can be defined over $\mathbf{F}_3$. Then since $\mathrm{PGL}_2(\mathbf{F}_3) \simeq S_4$ every such representation is modular by the theorem of Langlands and Tunnell mentioned above. In particular, every representation into $\mathrm{GL}_2(\mathbf{Z}_3)$ whose reduction satisfies the given conditions is modular. We deduce:

THEOREM 0.3. *Suppose that E is an elliptic curve defined over $\mathbf{Q}$ and that ρ_0 is the Galois action on the 3-division points. Suppose that E has the following properties:*

(i) *E has good or multiplicative reduction at 3.*

(ii) *ρ_0 is absolutely irreducible when restricted to $\mathbf{Q}\left(\sqrt{-3}\right)$.*

(iii) *For any $q \equiv -1 \bmod 3$ either $\rho_0|_{D_q}$ is reducible over the algebraic closure or $\rho_0|_{I_q}$ is absolutely irreducible.*

Then E is modular.

We should point out that while the properties of the zeta function follow directly from Theorem 0.2 the stronger version that E is covered by $X_0(N)$

requires also the isogeny theorem proved by Faltings (and earlier by Serre when E has nonintegral j-invariant, a case which includes the semistable curves). We note that if E is modular then so is any twist of E, so we could relax condition (i) somewhat.

The important class of semistable curves, i.e., those with square-free conductor, satisfies (i) and (iii) but not necessarily (ii). If (ii) fails then in fact ρ_0 is reducible. Rather surprisingly, Theorem 0.2 can often be applied in this case also by showing that the representation on the 5-division points also occurs for another elliptic curve which Theorem 0.3 has already proved modular. Thus Theorem 0.2 is applied this time with $p = 5$. This argument, which is explained in Chapter 5, is the only part of the paper which really uses deformations of the elliptic curve rather than deformations of the Galois representation. The argument works more generally than in the semistable case but in this setting we obtain the following theorem:

THEOREM 0.4. *Suppose that E is a semistable elliptic curve defined over $\mathbf{Q}$. Then E is modular.*

More general families of elliptic curves which are modular are given in Chapter 5.

In 1986, stimulated by an ingenious idea of Frey [Fr], Serre conjectured and Ribet proved (in [Ri1]) a property of the Galois representations associated to modular forms which enabled Ribet to show that Theorem 0.4 implies 'Fermat's Last Theorem'. Frey's suggestion, in the notation of the following theorem, was to show that the (hypothetical) elliptic curve $y^2 = x(x+u^p)(x-v^p)$ could not be modular. Such elliptic curves had already been studied in [He] but without the connection with modular forms. Serre made precise the idea of Frey by proposing a conjecture on modular forms which meant that the representation on the p-division points of this particular elliptic curve, if modular, would be associated to a form of conductor 2. This, by a simple inspection, could not exist. Serre's conjecture was then proved by Ribet in the summer of 1986. However, one still needed to know that the curve in question would have to be modular, and this is accomplished by Theorem 0.4. We have then (finally!):

THEOREM 0.5. *Suppose that $u^p + v^p + w^p = 0$ with $u, v, w \in \mathbf{Q}$ and $p \geq 3$, then $uvw = 0$.*

The second result we prove about the conjecture does not require the assumption that ρ_0 be modular (since it is already known in this case).

THEOREM 0.6. *Suppose that ρ_0 is irreducible and satisfies the hypotheses of the conjecture, including* (I) *above. Suppose further that*

(i) $\rho_0 = \mathrm{Ind}_L^{\mathbf{Q}}\, \kappa_0$ *for a character κ_0 of an imaginary quadratic extension L of $\mathbf{Q}$ which is unramified at p.*

(ii) $\det \rho_0|_{I_p} = \omega$.

Then a representation ρ as in the conjecture does indeed come from a modular form.

This theorem can also be used to prove that certain families of elliptic curves are modular. In this summary we have only described the principal theorems associated to Galois representations and elliptic curves. Our results concerning generalized class groups are described in Theorem 3.3.

The following is an account of the origins of this work and of the more specialized developments of the 1980's that affected it. I began working on these problems in the late summer of 1986 immediately on learning of Ribet's result. For several years I had been working on the Iwasawa conjecture for totally real fields and some applications of it. In the process, I had been using and developing results on ℓ-adic representations associated to Hilbert modular forms. It was therefore natural for me to consider the problem of modularity from the point of view of ℓ-adic representations. I began with the assumption that the reduction of a given ordinary ℓ-adic representation was reducible and tried to prove under this hypothesis that the representation itself would have to be modular. I hoped rather naively that in this situation I could apply the techniques of Iwasawa theory. Even more optimistically I hoped that the case $\ell = 2$ would be tractable as this would suffice for the study of the curves used by Frey. From now on and in the main text, we write p for ℓ because of the connections with Iwasawa theory.

After several months studying the 2-adic representation, I made the first real breakthrough in realizing that I could use the 3-adic representation instead: the Langlands-Tunnell theorem meant that ρ_3, the mod 3 representation of any given elliptic curve over $\mathbf{Q}$, would necessarily be modular. This enabled me to try inductively to prove that the $\mathrm{GL}_2(\mathbf{Z}/3^n\mathbf{Z})$ representation would be modular for each n. At this time I considered only the ordinary case. This led quickly to the study of $H^i(\mathrm{Gal}(F_\infty/\mathbf{Q}), W_f)$ for $i = 1$ and 2, where F_∞ is the splitting field of the $\mathfrak{m}$-adic torsion on the Jacobian of a suitable modular curve, $\mathfrak{m}$ being the maximal ideal of a Hecke ring associated to ρ_3 and W_f the module associated to a modular form f described in Chapter 1. More specifically, I needed to compare this cohomology with the cohomology of $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$ acting on the same module.

I tried to apply some ideas from Iwasawa theory to this problem. In my solution to the Iwasawa conjecture for totally real fields [Wi4], I had introduced

a new technique in order to deal with the trivial zeroes. It involved replacing the standard Iwasawa theory method of considering the fields in the cyclotomic $\mathbf{Z}_p$-extension by a similar analysis based on a choice of infinitely many distinct primes $q_i \equiv 1 \bmod p^{n_i}$ with $n_i \to \infty$ as $i \to \infty$. Some aspects of this method suggested that an alternative to the standard technique of Iwasawa theory, which seemed problematic in the study of W_f, might be to make a comparison between the cohomology groups as Σ varies but with the field $\mathbf{Q}$ fixed. The new principle said roughly that the unramified cohomology classes are trapped by the tamely ramified ones. After reading the paper [Gre1], I realized that the duality theorems in Galois cohomology of Poitou and Tate would be useful for this. The crucial extract from this latter theory is in Section 2 of Chapter 1.

In order to put these ideas into practice I developed in a naive form the techniques of the first two sections of Chapter 2. This drew in particular on a detailed study of all the congruences between f and other modular forms of differing levels, a theory that had been initiated by Hida and Ribet. The outcome was that I could estimate the first cohomology group well under two assumptions, first that a certain subgroup of the second cohomology group vanished and second that the form f was chosen at the minimal level for $\mathfrak{m}$. These assumptions were much too restrictive to be really effective but at least they pointed in the right direction. Some of these arguments are to be found in the second section of Chapter 1 and some form the first weak approximation to the argument in Chapter 3. At that time, however, I used auxiliary primes $q \equiv -1 \bmod p$ when varying Σ as the geometric techniques I worked with did not apply in general for primes $q \equiv 1 \bmod p$. (This was for much the same reason that the reduction of level argument in [Ri1] is much more difficult when $q \equiv 1 \bmod p$.) In all this work I used the more general assumption that ρ_p was modular rather than the assumption that $p = 3$.

In the late 1980's, I translated these ideas into ring-theoretic language. A few years previously Hida had constructed some explicit one-parameter families of Galois representations. In an attempt to understand this, Mazur had been developing the language of deformations of Galois representations. Moreover, Mazur realized that the universal deformation rings he found should be given by Hecke rings, at least in certain special cases. This critical conjecture refined the expectation that all ordinary liftings of modular representations should be modular. In making the translation to this ring-theoretic language I realized that the vanishing assumption on the subgroup of H^2 which I had needed should be replaced by the stronger condition that the Hecke rings were complete intersections. This fitted well with their being deformation rings where one could estimate the number of generators and relations and so made the original assumption more plausible.

To be of use, the deformation theory required some development. Apart from some special examples examined by Boston and Mazur there had been

little work on it. I checked that one could make the appropriate adjustments to the theory in order to describe deformation theories at the minimal level. In the fall of 1989, I set Ramakrishna, then a student of mine at Princeton, the task of proving the existence of a deformation theory associated to representations arising from finite flat group schemes over $\mathbf{Z}_p$. This was needed in order to remove the restriction to the ordinary case. These developments are described in the first section of Chapter 1 although the work of Ramakrishna was not completed until the fall of 1991. For a long time the ring-theoretic version of the problem, although more natural, did not look any simpler. The usual methods of Iwasawa theory when translated into the ring-theoretic language seemed to require unknown principles of base change. One needed to know the exact relations between the Hecke rings for different fields in the cyclotomic $\mathbf{Z}_p$-extension of $\mathbf{Q}$, and not just the relations up to torsion.

The turning point in this and indeed in the whole proof came in the spring of 1991. In searching for a clue from commutative algebra I had been particularly struck some years earlier by a paper of Kunz [Ku2]. I had already needed to verify that the Hecke rings were Gorenstein in order to compute the congruences developed in Chapter 2. This property had first been proved by Mazur in the case of prime level and his argument had already been extended by other authors as the need arose. Kunz's paper suggested the use of an invariant (the η-invariant of the appendix) which I saw could be used to test for isomorphisms between Gorenstein rings. A different invariant (the $\mathfrak{p}/\mathfrak{p}^2$-invariant of the appendix) I had already observed could be used to test for isomorphisms between complete intersections. It was only on reading Section 6 of [Ti2] that I learned that it followed from Tate's account of Grothendieck duality theory for complete intersections that these two invariants were equal for such rings. Not long afterwards I realized that, unlikely though it seemed at first, the equality of these invariants was actually a criterion for a Gorenstein ring to be a complete intersection. These arguments are given in the appendix.

The impact of this result on the main problem was enormous. Firstly, the relationship between the Hecke rings and the deformation rings could be tested just using these two invariants. In particular I could provide the inductive argument of Section 3 of Chapter 2 to show that if all liftings with restricted ramification are modular then all liftings are modular. This I had been trying to do for a long time but without success until the breakthrough in commutative algebra. Secondly, by means of a calculation of Hida summarized in [Hi2] the main problem could be transformed into a problem about class numbers of a type well-known in Iwasawa theory. In particular, I could check this in the ordinary CM case using the recent theorems of Rubin and Kolyvagin. This is the content of Chapter 4. Thirdly, it meant that for the first time it could be verified that infinitely many j-invariants were modular. Finally, it meant that I could focus on the minimal level where the estimates given by my earlier

Galois cohomology calculations looked more promising. Here I was also using the work of Ribet and others on Serre's conjecture (the same work of Ribet that had linked Fermat's Last Theorem to modular forms in the first place) to know that there was a minimal level.

The class number problem was of a type well-known in Iwasawa theory and in the ordinary case had already been conjectured by Coates and Schmidt. However, the traditional methods of Iwasawa theory did not seem quite sufficient in this case and, as explained earlier, when translated into the ring-theoretic language seemed to require unknown principles of base change. So instead I developed further the idea of using auxiliary primes to replace the change of field that is used in Iwasawa theory. The Galois cohomology estimates described in Chapter 3 were now much stronger, although at that time I was still using primes $q \equiv -1 \bmod p$ for the argument. The main difficulty was that although I knew how the η-invariant changed as one passed to an auxiliary level from the results of Chapter 2, I did not know how to estimate the change in the $\mathfrak{p}/\mathfrak{p}^2$-invariant precisely. However, the method did give the right bound for the generalised class group, or Selmer group as it is often called in this context, under the additional assumption that the minimal Hecke ring was a complete intersection.

I had earlier realized that ideally what I needed in this method of auxiliary primes was a replacement for the power series ring construction one obtains in the more natural approach based on Iwasawa theory. In this more usual setting, the projective limit of the Hecke rings for the varying fields in a cyclotomic tower would be expected to be a power series ring, at least if one assumed the vanishing of the μ-invariant. However, in the setting with auxiliary primes where one would change the level but not the field, the natural limiting process did not appear to be helpful, with the exception of the closely related and very important construction of Hida [Hi1]. This method of Hida often gave one step towards a power series ring in the ordinary case. There were also tenuous hints of a patching argument in Iwasawa theory ([Scho], [Wi4, §10]), but I searched without success for the key.

Then, in August, 1991, I learned of a new construction of Flach [Fl] and quickly became convinced that an extension of his method was more plausible. Flach's approach seemed to be the first step towards the construction of an Euler system, an approach which would give the precise upper bound for the size of the Selmer group if it could be completed. By the fall of 1992, I believed I had achieved this and began then to consider the remaining case where the mod 3 representation was assumed reducible. For several months I tried simply to repeat the methods using deformation rings and Hecke rings. Then unexpectedly in May 1993, on reading of a construction of twisted forms of modular curves in a paper of Mazur [Ma3], I made a crucial and surprising breakthrough: I found the argument using families of elliptic curves with a

common ρ_5 which is given in Chapter 5. Believing now that the proof was complete, I sketched the whole theory in three lectures in Cambridge, England on June 21–23. However, it became clear to me in the fall of 1993 that the construction of the Euler system used to extend Flach's method was incomplete and possibly flawed.

Chapter 3 follows the original approach I had taken to the problem of bounding the Selmer group but had abandoned on learning of Flach's paper. Darmon encouraged me in February, 1994, to explain the reduction to the complete intersection property, as it gave a quick way to exhibit infinite families of modular j-invariants. In presenting it in a lecture at Princeton, I made, almost unconsciously, a critical switch to the special primes used in Chapter 3 as auxiliary primes. I had only observed the existence and importance of these primes in the fall of 1992 while trying to extend Flach's work. Previously, I had only used primes $q \equiv -1 \bmod p$ as auxiliary primes. In hindsight this change was crucial because of a development due to de Shalit. As explained before, I had realized earlier that Hida's theory often provided one step towards a power series ring at least in the ordinary case. At the Cambridge conference de Shalit had explained to me that for primes $q \equiv 1 \bmod p$ he had obtained a version of Hida's results. But except for explaining the complete intersection argument in the lecture at Princeton, I still did not give any thought to my initial approach, which I had put aside since the summer of 1991, since I continued to believe that the Euler system approach was the correct one.

Meanwhile in January, 1994, R. Taylor had joined me in the attempt to repair the Euler system argument. Then in the spring of 1994, frustrated in the efforts to repair the Euler system argument, I began to work with Taylor on an attempt to devise a new argument using $p = 2$. The attempt to use $p = 2$ reached an impasse at the end of August. As Taylor was still not convinced that the Euler system argument was irreparable, I decided in September to take one last look at my attempt to generalise Flach, if only to formulate more precisely the obstruction. In doing this I came suddenly to a marvelous revelation: I saw in a flash on September 19th, 1994, that de Shalit's theory, if generalised, could be used together with duality to glue the Hecke rings at suitable auxiliary levels into a power series ring. I had unexpectedly found the missing key to my old abandoned approach. It was the old idea of picking q_i's with $q_i \equiv 1 \bmod p^{n_i}$ and $n_i \to \infty$ as $i \to \infty$ that I used to achieve the limiting process. The switch to the special primes of Chapter 3 had made all this possible.

After I communicated the argument to Taylor, we spent the next few days making sure of the details. The full argument, together with the deduction of the complete intersection property, is given in [TW].

In conclusion the key breakthrough in the proof had been the realization in the spring of 1991 that the two invariants introduced in the appendix could be used to relate the deformation rings and the Hecke rings. In effect the η-

invariant could be used to count Galois representations. The last step after the June, 1993, announcement, though elusive, was but the conclusion of a long process whose purpose was to replace, in the ring-theoretic setting, the methods based on Iwasawa theory by methods based on the use of auxiliary primes.

One improvement that I have not included but which might be used to simplify some of Chapter 2 is the observation of Lenstra that the criterion for Gorenstein rings to be complete intersections can be extended to more general rings which are finite and free as $\mathbf{Z}_p$-modules. Faltings has pointed out an improvement, also not included, which simplifies the argument in Chapter 3 and [TW]. This is however explained in the appendix to [TW].

It is a pleasure to thank those who read carefully a first draft of some of this paper after the Cambridge conference and particularly N. Katz who patiently answered many questions in the course of my work on Euler systems, and together with Illusie read critically the Euler system argument. Their questions led to my discovery of the problem with it. Katz also listened critically to my first attempts to correct it in the fall of 1993. I am grateful also to Taylor for his assistance in analyzing in depth the Euler system argument. I am indebted to F. Diamond for his generous assistance in the preparation of the final version of this paper. In addition to his many valuable suggestions, several others also made helpful comments and suggestions especially Conrad, de Shalit, Faltings, Ribet, Rubin, Skinner and Taylor. Finally, I am most grateful to H. Darmon for his encouragement to reconsider my old argument. Although I paid no heed to his advice at the time, it surely left its mark.

Table of Contents

Chapter 1	1.	Deformations of Galois representations
	2.	Some computations of cohomology groups
	3.	Some results on subgroups of $\mathrm{GL}_2(k)$
Chapter 2	1.	The Gorenstein property
	2.	Congruences between Hecke rings
	3.	The main conjectures
Chapter 3		Estimates for the Selmer group
Chapter 4	1.	The ordinary CM case
	2.	Calculation of η
Chapter 5		Application to elliptic curves
Appendix		
References		

Chapter 1

This chapter is devoted to the study of certain Galois representations. In the first section we introduce and study Mazur's deformation theory and discuss various refinements of it. These refinements will be needed later to make precise the correspondence between the universal deformation rings and the Hecke rings in Chapter 2. The main results needed are Proposition 1.2 which is used to interpret various generalized cotangent spaces as Selmer groups and (1.7) which later will be used to study them. At the end of the section we relate these Selmer groups to ones used in the Bloch-Kato conjecture, but this connection is not needed for the proofs of our main results.

In the second section we extract from the results of Poitou and Tate on Galois cohomology certain general relations between Selmer groups as Σ varies, as well as between Selmer groups and their duals. The most important observation of the third section is Lemma 1.10(i) which guarantees the existence of the special primes used in Chapter 3 and [TW].

1. Deformations of Galois representations

Let p be an odd prime. Let Σ be a finite set of primes including p and let $\mathbf{Q}_\Sigma$ be the maximal extension of $\mathbf{Q}$ unramified outside this set and ∞. Throughout we fix an embedding of $\overline{\mathbf{Q}}$, and so also of $\mathbf{Q}_\Sigma$, in $\mathbf{C}$. We will also fix a choice of decomposition group D_q for all primes q in $\mathbf{Z}$. Suppose that k is a finite field of characteristic p and that

$$\rho_0\colon \operatorname{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \operatorname{GL}_2(k) \tag{1.1}$$

is an irreducible representation. In contrast to the introduction we will assume in the rest of the paper that ρ_0 comes with its field of definition k. Suppose further that $\det\rho_0$ is odd. In particular this implies that the smallest field of definition for ρ_0 is given by the field k_0 generated by the traces but we will not assume that $k = k_0$. It also implies that ρ_0 is absolutely irreducible. We consider the deformations $[\rho]$ to $\operatorname{GL}_2(A)$ of ρ_0 in the sense of Mazur [Ma1]. Thus if $W(k)$ is the ring of Witt vectors of k, A is to be a complete Noetherian local $W(k)$-algebra with residue field k and maximal ideal m, and a deformation $[\rho]$ is just a strict equivalence class of homomorphisms $\rho\colon \operatorname{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \operatorname{GL}_2(A)$ such that $\rho \bmod m = \rho_0$, two such homomorphisms being called strictly equivalent if one can be brought to the other by conjugation by an element of $\ker : \operatorname{GL}_2(A) \to \operatorname{GL}_2(k)$. We often simply write ρ instead of $[\rho]$ for the equivalence class.

We will restrict our choice of ρ_0 further by assuming that either:

(i) ρ_0 is *ordinary*; viz., the restriction of ρ_0 to the decomposition group D_p has (for a suitable choice of basis) the form

$$\rho_0\Big|_{D_p} \approx \begin{pmatrix} \chi_1 & * \\ 0 & \chi_2 \end{pmatrix} \tag{1.2}$$

where χ_1 and χ_2 are homomorphisms from D_p to k^* with χ_2 unramified. Moreover we require that $\chi_1 \neq \chi_2$. We do allow here that $\rho_0|_{D_p}$ be semisimple. (If χ_1 and χ_2 are both unramified and $\rho_0|_{D_p}$ is semisimple then we fix our choices of χ_1 and χ_2 once and for all.)

(ii) ρ_0 is *flat* at p but not ordinary (cf. [Se1] where the terminology *finite* is used); viz., $\rho_0|_{D_p}$ is the representation associated to a finite flat group scheme over $\mathbf{Z}_p$ but is not ordinary in the sense of (i). (In general when we refer to the flat case we will mean that ρ_0 is assumed not to be ordinary unless we specify otherwise.) We will assume also that $\det \rho_0|_{I_p} = \omega$ where I_p is an inertia group at p and ω is the Teichmüller character giving the action on p^{th} roots of unity.

In case (ii) it follows from results of Raynaud that $\rho_0|_{D_p}$ is absolutely irreducible and one can describe $\rho_0|_{I_p}$ explicitly. For extending a Jordan-Hölder series for the representation space (as an I_p-module) to one for finite flat group schemes (cf. [Ray1]) we observe first that the trivial character does not occur on a subquotient, as otherwise (using the classification of Oort-Tate or Raynaud) the group scheme would be ordinary. So we find by Raynaud's results, that $\rho_0|_{I_p} \underset{k}{\otimes} \bar{k} \simeq \psi_1 \oplus \psi_2$ where ψ_1 and ψ_2 are the two fundamental characters of degree 2 (cf. Corollary 3.4.4 of [Ray1]). Since ψ_1 and ψ_2 do not extend to characters of $\operatorname{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$, $\rho_0|_{D_p}$ must be absolutely irreducible.

We will sometimes wish to make one of the following restrictions on the deformations we allow:

(i) (a) *Selmer deformations.* In this case we assume that ρ_0 is ordinary, with notation as above, and that the deformation has a representative $\rho : \operatorname{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \operatorname{GL}_2(A)$ with the property that (for a suitable choice of basis)

$$\rho|_{D_p} \approx \begin{pmatrix} \tilde{\chi}_1 & * \\ 0 & \tilde{\chi}_2 \end{pmatrix}$$

with $\tilde{\chi}_2$ unramified, $\tilde{\chi}_2 \equiv \chi_2 \bmod m$, and $\det \rho|_{I_p} = \varepsilon\omega^{-1}\chi_1\chi_2$ where ε is the cyclotomic character, ε: $\operatorname{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \mathbf{Z}_p^*$, giving the action on all p-power roots of unity, ω is of order prime to p satisfying $\omega \equiv \varepsilon \bmod p$, and χ_1 and χ_2 are the characters of (i) viewed as taking values in $k^* \hookrightarrow A^*$.

(i) (b) *Ordinary deformations.* The same as in (i)(a) but with no condition on the determinant.

(i) (c) *Strict deformations.* This is a variant on (i) (a) which we only use when $\rho_0|_{D_p}$ is not semisimple and not flat (i.e. not associated to a finite flat group scheme). We also assume that $\chi_1\chi_2^{-1} = \omega$ in this case. Then a strict deformation is as in (i)(a) except that we assume in addition that $(\tilde{\chi}_1/\tilde{\chi}_2)|_{D_p} = \varepsilon$.

(ii) *Flat (at p) deformations.* We assume that each deformation ρ to $\mathrm{GL}_2(A)$ has the property that for any quotient $A/\mathfrak{a}$ of finite order $\rho|_{D_p} \bmod \mathfrak{a}$ is the Galois representation associated to the $\bar{\mathbf{Q}}_p$-points of a finite flat group scheme over $\mathbf{Z}_p$.

In each of these four cases, as well as in the unrestricted case (in which we impose no local restriction at p) one can verify that Mazur's use of Schlessinger's criteria [Sch] proves the existence of a universal deformation

$$\rho\colon \mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \rightarrow \mathrm{GL}_2(R).$$

In the ordinary and unrestricted case this was proved by Mazur and in the flat case by Ramakrishna [Ram]. The other cases require minor modifications of Mazur's argument. We denote the universal ring R_Σ in the unrestricted case and $R_\Sigma^{\mathrm{se}}, R_\Sigma^{\mathrm{ord}}, R_\Sigma^{\mathrm{str}}, R_\Sigma^{\mathrm{f}}$ in the other four cases. We often omit the Σ if the context makes it clear.

There are certain generalizations to all of the above which we will also need. The first is that instead of considering $W(k)$ -algebras A we may consider $\mathcal{O}$-algebras for $\mathcal{O}$ the ring of integers of any local field with residue field k. If we need to record which $\mathcal{O}$ we are using we will write $R_{\Sigma,\mathcal{O}}$ etc. It is easy to see that the natural local map of local $\mathcal{O}$-algebras

$$R_{\Sigma,\mathcal{O}} \rightarrow R_\Sigma \underset{W(k)}{\otimes} \mathcal{O}$$

is an isomorphism because for functorial reasons the map has a natural section which induces an isomorphism on Zariski tangent spaces at closed points, and one can then use Nakayama's lemma. Note, however, that if we change the residue field via $i : k \hookrightarrow k'$ then we have a new deformation problem associated to the representation $\rho_0' = i \circ \rho_0$. There is again a natural map of $W(k')$-algebras

$$R(\rho_0') \rightarrow R \underset{W(k)}{\otimes} W(k')$$

which is an isomorphism on Zariski tangent spaces. One can check that this is again an isomorphism by considering the subring R_1 of $R(\rho_0')$ defined as the subring of all elements whose reduction modulo the maximal ideal lies in k. Since $R(\rho_0')$ is a finite R_1-module, R_1 is also a complete local Noetherian ring

with residue field k. The universal representation associated to ρ_0' is defined over R_1 and the universal property of R then defines a map $R \to R_1$. So we obtain a section to the map $R(\rho_0') \to R \underset{W(k)}{\otimes} W(k')$ and the map is therefore an isomorphism. (I am grateful to Faltings for this observation.) We will also need to extend the consideration of $\mathcal{O}$-algebras to the restricted cases. In each case we can require A to be an $\mathcal{O}$-algebra and again it is easy to see that $R^{\cdot}_{\Sigma,\mathcal{O}} \simeq R^{\cdot}_{\Sigma} \underset{W(k)}{\otimes} \mathcal{O}$ in each case.

The second generalization concerns primes $q \neq p$ which are ramified in ρ_0. We distinguish three special cases (types (A) and (C) need not be disjoint):

(A) $\rho_0|_{D_q} = \left(\begin{smallmatrix} \chi_1 & * \\ & \chi_2 \end{smallmatrix}\right)$ for a suitable choice of basis, with χ_1 and χ_2 unramified, $\chi_1 \chi_2^{-1} = \omega$ and the fixed space of I_q of dimension 1,

(B) $\rho_0|_{I_q} = \left(\begin{smallmatrix} \chi_q & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\chi_q \neq 1$, for a suitable choice of basis,

(C) $H^1(\mathbf{Q}_q, W_\lambda) = 0$ where W_λ is as defined in (1.6).

Then in each case we can define a suitable deformation theory by imposing additional restrictions on those we have already considered, namely:

(A) $\rho|_{D_q} = \left(\begin{smallmatrix} \psi_1 & * \\ & \psi_2 \end{smallmatrix}\right)$ for a suitable choice of basis of A^2 with ψ_1 and ψ_2 unramified and $\psi_1 \psi_2^{-1} = \varepsilon$;

(B) $\rho|_{I_q} = \left(\begin{smallmatrix} \chi_q & 0 \\ 0 & 1 \end{smallmatrix}\right)$ for a suitable choice of basis (χ_q of order prime to p, so the *same* character as above);

(C) $\det \rho|_{I_q} = \det \rho_0|_{I_q}$, i.e., of order prime to p.

Thus if $\mathcal{M}$ is a set of primes in Σ distinct from p and each satisfying one of (A), (B) or (C) for ρ_0, we will impose the corresponding restriction at each prime in $\mathcal{M}$.

Thus to each set of data $\mathcal{D} = \{\cdot, \Sigma, \mathcal{O}, \mathcal{M}\}$ where $\cdot$ is Se, str, ord, flat or unrestricted, we can associate a deformation theory to ρ_0 provided

$$\rho_0\colon \operatorname{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \operatorname{GL}_2(k) \tag{1.3}$$

is itself of type $\mathcal{D}$ and $\mathcal{O}$ is the ring of integers of a totally ramified extension of $W(k)$; ρ_0 is ordinary if $\cdot$ is Se or ord, strict if $\cdot$ is strict and flat if $\cdot$ is fl (meaning flat); ρ_0 is of type $\mathcal{M}$, i.e., of type (A), (B) or (C) at each ramified prime $q \neq p$, $q \in \mathcal{M}$. We allow different types at different q's. We will refer to these as the standard deformation theories and write $R_\mathcal{D}$ for the universal ring associated to $\mathcal{D}$ and $\rho_\mathcal{D}$ for the universal deformation (or even ρ if $\mathcal{D}$ is clear from the context).

We note here that if $\mathcal{D} = (\text{ord}, \Sigma, \mathcal{O}, \mathcal{M})$ and $\mathcal{D}' = (\text{Se}, \Sigma, \mathcal{O}, \mathcal{M})$ then there is a simple relation between $R_\mathcal{D}$ and $R_{\mathcal{D}'}$. Indeed there is a natural map

$R_{\mathcal{D}} \to R_{\mathcal{D}'}$ by the universal property of $R_{\mathcal{D}}$, and its kernel is a principal ideal generated by $T = \varepsilon^{-1}(\gamma)\det\rho_{\mathcal{D}}(\gamma) - 1$ where $\gamma \in \mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$ is any element whose restriction to $\mathrm{Gal}(\mathbf{Q}_\infty/\mathbf{Q})$ is a generator (where $\mathbf{Q}_\infty$ is the $\mathbf{Z}_p$-extension of $\mathbf{Q}$) and whose restriction to $\mathrm{Gal}(\mathbf{Q}(\zeta_{Np})/\mathbf{Q})$ is trivial for any N prime to p with $\zeta_N \in \mathbf{Q}_\Sigma$, ζ_N being a primitive N^{th} root of 1:

$$R_{\mathcal{D}}/T \simeq R_{\mathcal{D}'}. \tag{1.4}$$

It turns out that under the hypothesis that ρ_0 is strict, i.e. that $\rho_0|_{D_p}$ is not associated to a finite flat group scheme, the deformation problems in (i)(a) and (i)(c) are the same; i.e., every Selmer deformation is already a strict deformation. This was observed by Diamond. The argument is local, so the decomposition group D_p could be replaced by $\mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$.

PROPOSITION 1.1 (Diamond). *Suppose that $\pi{:}D_p \to \mathrm{GL}_2(A)$ is a continuous representation where A is an Artinian local ring with residue field k, a finite field of characteristic p. Suppose $\pi \approx \left(\begin{smallmatrix}\chi_1\varepsilon & * \\ 0 & \chi_2\end{smallmatrix}\right)$ with χ_1 and χ_2 unramified and $\chi_1 \neq \chi_2$. Then the residual representation $\bar{\pi}$ is associated to a finite flat group scheme over $\mathbf{Z}_p$.*

Proof (taken from [Dia, Prop. 6.1]). We may replace π by $\pi \otimes \chi_2^{-1}$ and we let $\varphi = \chi_1\chi_2^{-1}$. Then $\pi \cong \left(\begin{smallmatrix}\varphi\varepsilon & t \\ 0 & 1\end{smallmatrix}\right)$ determines a cocycle $t : D_p \to M(1)$ where M is a free A-module of rank one on which D_p acts via φ. Let u denote the cohomology class in $H^1(D_p, M(1))$ defined by t, and let u_0 denote its image in $H^1(D_p, M_0(1))$ where $M_0 = M/\mathfrak{m}M$. Let $G = \ker\varphi$ and let F be the fixed field of G (so F is a finite unramified extension of $\mathbf{Q}_p$). Choose n so that $p^nA = 0$. Since $H^2(G, \mu_{p^r}) \to H^2(G, \mu_{p^s})$ is injective for $r \le s$, we see that the natural map of $A[D_p/G]$-modules $H^1(G, \mu_{p^n}) \otimes_{\mathbf{Z}_p} M \to H^1(G, M(1))$ is an isomorphism. By Kummer theory, we have $H^1(G, M(1)) \cong F^\times/(F^\times)^{p^n} \otimes_{\mathbf{Z}_p} M$ as D_p-modules. Now consider the commutative diagram

$$\begin{array}{ccccc}
H^1(G, M(1))^{D_p} & \xrightarrow{\sim} & ((F^\times/(F^\times)^{p^n} \otimes_{\mathbf{Z}_p} M)^{D_p} & \longrightarrow & M^{D_p} \\
\downarrow & & \downarrow & & \downarrow \\
H^1(G, M_0(1)) & \xrightarrow{\sim} & (F^\times/(F^\times)^p) \otimes_{\mathbf{F}_p} M_0 & \longrightarrow & M_0
\end{array},$$

where the right-hand horizontal maps are induced by $v_p : F^\times \to \mathbf{Z}$. If $\varphi \neq 1$, then $M^{D_p} \subset \mathfrak{m}M$, so that the element $\mathrm{res}\, u_0$ of $H^1(G, M_0(1))$ is in the image of $(\mathcal{O}_F^\times/(\mathcal{O}_F^\times)^p) \otimes_{\mathbf{F}_p} M_0$. But this means that $\bar{\pi}$ is "peu ramifié" in the sense of [Se] and therefore $\bar{\pi}$ comes from a finite flat group scheme. (See [E1, (8.2)].)

Remark. Diamond also observes that essentially the same proof shows that if $\pi : \mathrm{Gal}(\bar{\mathbf{Q}}_q/\mathbf{Q}_q) \to \mathrm{GL}_2(A)$, where A is a complete local Noetherian

ring with residue field k, has the form $\pi|_{I_q} \cong \left(\begin{smallmatrix}1 & *\\ 0 & 1\end{smallmatrix}\right)$ with $\bar{\pi}$ ramified then π is of type (A).

Globally, Proposition 1.1 says that if ρ_0 is strict and if $\mathcal{D} = (\text{Se}, \Sigma, \mathcal{O}, \mathcal{M})$ and $\mathcal{D}' = (\text{str}, \Sigma, \mathcal{O}, \mathcal{M})$ then the natural map $R_{\mathcal{D}} \to R_{\mathcal{D}'}$ is an isomorphism.

In each case the tangent space of $R_{\mathcal{D}}$ may be computed as in [Ma1]. Let λ be a uniformizer for $\mathcal{O}$ and let $U_\lambda \simeq k^2$ be the representation space for ρ_0. (The motivation for the subscript λ will become apparent later.) Let V_λ be the representation space of $\text{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$ on $\text{Ad}\,\rho_0 = \text{Hom}_k(U_\lambda, U_\lambda) \simeq M_2(k)$. Then there is an isomorphism of k-vector spaces (cf. the proof of Prop. 1.2 below)

$$\text{(1.5)} \qquad \text{Hom}_k(m_{\mathcal{D}}/(m_{\mathcal{D}}^2, \lambda), k) \simeq H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_\lambda)$$

where $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_\lambda)$ is a subspace of $H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, V_\lambda)$ which we now describe and $m_{\mathcal{D}}$ is the maximal ideal of $R_{\mathcal{D}}$. It consists of the cohomology classes which satisfy certain local restrictions at p and at the primes in $\mathcal{M}$. We call $m_{\mathcal{D}}/(m_{\mathcal{D}}^2, \lambda)$ the reduced cotangent space of $R_{\mathcal{D}}$.

We begin with p. First we may write (since $p \neq 2$), as $k[\text{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})]$-modules,

$$\text{(1.6)} \quad V_\lambda = W_\lambda \oplus k, \text{ where } W_\lambda = \{f \in \text{Hom}_k(U_\lambda, U_\lambda)\colon \text{trace}\, f = 0\} \simeq (\text{Sym}^2 \otimes \det^{-1})\rho_0$$

and k is the one-dimensional subspace of scalar multiplications. Then if ρ_0 is ordinary the action of D_p on U_λ induces a filtration of U_λ and also on W_λ and V_λ. Suppose we write these $0 \subset U_\lambda^0 \subset U_\lambda$, $0 \subset W_\lambda^0 \subset W_\lambda^1 \subset W_\lambda$ and $0 \subset V_\lambda^0 \subset V_\lambda^1 \subset V_\lambda$. Thus U_λ^0 is defined by the requirement that D_p act on it via the character χ_1 (cf. (1.2)) and on U_λ/U_λ^0 via χ_2. For W_λ the filtrations are defined by

$$W_\lambda^1 = \{f \in W_\lambda\colon\ f(U_\lambda^0) \subset U_\lambda^0\},$$
$$W_\lambda^0 = \{f \in W_\lambda^1\colon\ f = 0 \text{ on } U_\lambda^0\},$$

and the filtrations for V_λ are obtained by replacing W by V. We note that these filtrations are often characterized by the action of D_p. Thus the action of D_p on W_λ^0 is via χ_1/χ_2; on W_λ^1/W_λ^0 it is trivial and on W_λ/W_λ^1 it is via χ_2/χ_1. These determine the filtration if either χ_1/χ_2 is not quadratic or $\rho_0|_{D_p}$ is not semisimple. We define the k-vector spaces

$$V_\lambda^{\text{ord}} = \{f \in V_\lambda^1 : f = 0 \quad \text{in} \quad \text{Hom}(U_\lambda/U_\lambda^0, U_\lambda/U_\lambda^0)\},$$
$$H^1_{\text{Se}}(\mathbf{Q}_p, V_\lambda) = \ker\{H^1(\mathbf{Q}_p, V_\lambda) \to H^1(\mathbf{Q}_p^{\text{unr}}, V_\lambda/W_\lambda^0)\},$$
$$H^1_{\text{ord}}(\mathbf{Q}_p, V_\lambda) = \ker\{H^1(\mathbf{Q}_p, V_\lambda) \to H^1(\mathbf{Q}_p^{\text{unr}}, V_\lambda/V_\lambda^{\text{ord}})\},$$
$$H^1_{\text{str}}(\mathbf{Q}_p, V_\lambda) = \ker\{H^1(\mathbf{Q}_p, V_\lambda) \to H^1(\mathbf{Q}_p, W_\lambda/W_\lambda^0) \oplus H^1(\mathbf{Q}_p^{\text{unr}}, k)\}.$$

In the Selmer case we make an analogous definition for $H^1_{\mathrm{Se}}(\mathbf{Q}_p, W_\lambda)$ by replacing V_λ by W_λ, and similarly in the strict case. In the flat case we use the fact that there is a natural isomorphism of k-vector spaces

$$H^1(\mathbf{Q}_p, V_\lambda) \rightarrow \mathrm{Ext}^1_{k[D_p]}(U_\lambda, U_\lambda)$$

where the extensions are computed in the category of k-vector spaces with local Galois action. Then $H^1_{\mathrm{f}}(\mathbf{Q}_p, V_\lambda)$ is defined as the k-subspace of $H^1(\mathbf{Q}_p, V_\lambda)$ which is the inverse image of $\mathrm{Ext}^1_{\mathrm{fl}}(G, G)$, the group of extensions in the category of finite flat commutative group schemes over $\mathbf{Z}_p$ killed by p, G being the (unique) finite flat group scheme over $\mathbf{Z}_p$ associated to U_λ. By [Ray1] all such extensions in the inverse image even correspond to k-vector space schemes. For more details and calculations see [Ram].

For q different from p and $q \in \mathcal{M}$ we have three cases (A), (B), (C). In case (A) there is a filtration by D_q entirely analogous to the one for p. We write this $0 \subset W_\lambda^{0,q} \subset W_\lambda^{1,q} \subset W_\lambda$ and we set

$$H^1_{D_q}(\mathbf{Q}_q, V_\lambda) = \begin{cases} \ker : H^1(\mathbf{Q}_q, V_\lambda) \\ \qquad \rightarrow H^1(\mathbf{Q}_q, W_\lambda/W_\lambda^{0,q}) \oplus H^1(\mathbf{Q}_q^{\mathrm{unr}}, k) & \text{in case (A)} \\ \ker : H^1(\mathbf{Q}_q, V_\lambda) \\ \qquad \rightarrow H^1(\mathbf{Q}_q^{\mathrm{unr}}, V_\lambda) & \text{in case (B) or (C).} \end{cases}$$

Again we make an analogous definition for $H^1_{D_q}(\mathbf{Q}_q, W_\lambda)$ by replacing V_λ by W_λ and deleting the last term in case (A). We now define the k-vector space $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_\lambda)$ as

$$H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_\lambda) = \{\alpha \in H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, V_\lambda) : \quad \alpha_q \in H^1_{D_q}(\mathbf{Q}_q, V_\lambda) \text{ for all } q \in \mathcal{M}, \\ \alpha_p \in H^1_*(\mathbf{Q}_p, V_\lambda)\}$$

where $*$ is Se, str, ord, fl or unrestricted according to the type of $\mathcal{D}$. A similar definition applies to $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, W_\lambda)$ if $\cdot$ is Selmer or strict.

Now and for the rest of the section we are going to assume that ρ_0 arises from the reduction of the λ-adic representation associated to an eigenform. More precisely we assume that there is a normalized eigenform f of weight 2 and level N, divisible only by the primes in Σ, and that there is a prime λ of $\mathcal{O}_f$ such that $\rho_0 = \rho_{f,\lambda} \bmod \lambda$. Here $\mathcal{O}_f$ is the ring of integers of the field generated by the Fourier coefficients of f so the fields of definition of the two representations need not be the same. However we assume that $k \supseteq \mathcal{O}_{f,\lambda}/\lambda$ and we fix such an embedding so the comparison can be made over k. It will be convenient moreover to assume that if we are considering ρ_0 as being of type $\mathcal{D}$ then $\mathcal{D}$ is defined using $\mathcal{O}$-algebras where $\mathcal{O} \supseteq \mathcal{O}_{f,\lambda}$ is an unramified extension whose residue field is k. (Although this condition is unnecessary, it is convenient to use λ as the uniformizer for $\mathcal{O}$.) Finally we assume that $\rho_{f,\lambda}$

itself is of type $\mathcal{D}$. Again this is a slight abuse of terminology as we are really considering the extension of scalars $\rho_{f,\lambda} \underset{\mathcal{O}_{f,\lambda}}{\otimes} \mathcal{O}$ and not $\rho_{f,\lambda}$ itself, but we will do this without further mention if the context makes it clear. (The analysis of this section actually applies to any characteristic zero lifting of ρ_0 but in all our applications we will be in the more restrictive context we have described here.)

With these hypotheses there is a unique local homomorphism $R_{\mathcal{D}} \to \mathcal{O}$ of $\mathcal{O}$-algebras which takes the universal deformation to (the class of) $\rho_{f,\lambda}$. Let $\mathfrak{p}_{\mathcal{D}} = \ker : R_{\mathcal{D}} \to \mathcal{O}$. Let K be the field of fractions of $\mathcal{O}$ and let $U_f = (K/\mathcal{O})^2$ with the Galois action taken from $\rho_{f,\lambda}$. Similarly, let $V_f = \mathrm{Ad}\,\rho_{f,\lambda} \otimes_{\mathcal{O}} K/\mathcal{O} \simeq (K/\mathcal{O})^4$ with the adjoint representation so that

$$V_f \simeq W_f \oplus K/\mathcal{O}$$

where W_f has Galois action via $\mathrm{Sym}^2 \rho_{f,\lambda} \otimes \det \rho_{f,\lambda}^{-1}$ and the action on the second factor is trivial. Then if ρ_0 is ordinary the filtration of U_f under the $\mathrm{Ad}\,\rho$ action of D_p induces one on W_f which we write $0 \subset W_f^0 \subset W_f^1 \subset W_f$. Often to simplify the notation we will drop the index f from W_f^i, V_f etc. There is also a filtration on $W_{\lambda^n} = \{\ker\ \lambda^n\colon\ W_f \longrightarrow W_f\}$ given by $W_{\lambda^n}^i = W_{\lambda^n} \cap W^i$ (compatible with our previous description for $n = 1$). Likewise we write V_{λ^n} for $\{\ker\ \lambda^n\colon\ V_f \longrightarrow V_f\}$.

We now explain how to extend the definition of $H^1_{\mathcal{D}}$ to give meaning to $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ V_{\lambda^n})$ and $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ V)$ and these are $\mathcal{O}/\lambda^n$ and $\mathcal{O}$-modules, respectively. In the case where ρ_0 is ordinary the definitions are the same with V_{λ^n} or V replacing V_λ and $\mathcal{O}/\lambda^n$ or $K/\mathcal{O}$ replacing k. One checks easily that as $\mathcal{O}$-modules

$$H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n}) \simeq H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)_{\lambda^n}, \tag{1.7}$$

where as usual the subscript λ^n denotes the kernel of multiplication by λ^n. This just uses the divisibility of $H^0(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$ and $H^0(\mathbf{Q}_p,\ W/W^0)$ in the strict case. In the Selmer case one checks that for $m > n$ the kernel of

$$H^1(\mathbf{Q}_p^{\mathrm{unr}},\ V_{\lambda^n}/W_{\lambda^n}^0) \to H^1(\mathbf{Q}_p^{\mathrm{unr}},\ V_{\lambda^m}/W_{\lambda^m}^0)$$

has only the zero element fixed under $\mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}/\mathbf{Q}_p)$ and the ord case is similar. Checking conditions at $q \in \mathcal{M}$ is done with similar arguments. In the Selmer and strict cases we make analogous definitions with W_{λ^n} in place of V_{λ^n} and W in place of V and the analogue of (1.7) still holds.

We now consider the case where ρ_0 is flat (but not ordinary). We claim first that there is a natural map of $\mathcal{O}$-modules

$$H^1(\mathbf{Q}_p, V_{\lambda^n}) \to \mathrm{Ext}^1_{\mathcal{O}[D_p]}(U_{\lambda^m},\ U_{\lambda^n}) \tag{1.8}$$

for each $m \geq n$ where the extensions are of $\mathcal{O}$-modules with local Galois action. To describe this suppose that $\alpha \in H^1(\mathbf{Q}_p, V_{\lambda^n})$. Then we can associate to α a representation $\rho_\alpha\colon\ \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p) \to \mathrm{GL}_2(\mathcal{O}_n[\varepsilon])$ (where $\mathcal{O}_n[\varepsilon] =$

$\mathcal{O}[\varepsilon]/(\lambda^n\varepsilon, \varepsilon^2))$ which is an $\mathcal{O}$-algebra deformation of ρ_0 (see the proof of Proposition 1.1 below). Let $E = \mathcal{O}_n[\varepsilon]^2$ where the Galois action is via ρ_α. Then there is an exact sequence

$$\begin{array}{ccccccccc} 0 \longrightarrow & \varepsilon E/\lambda^m & \longrightarrow & E/\lambda^m & \longrightarrow & (E/\varepsilon)/\lambda^m & \longrightarrow 0 \\ & \wr\| & & & & \wr\| & \\ & U_{\lambda^n} & & & & U_{\lambda^m} & \end{array}$$

and hence an extension class in $\mathrm{Ext}^1(U_{\lambda^m}, U_{\lambda^n})$. One checks now that (1.8) is a map of $\mathcal{O}$-modules. We define $H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n})$ to be the inverse image of $\mathrm{Ext}^1_{\mathrm{fl}}(U_{\lambda^n}, U_{\lambda^n})$ under (1.8), i.e., those extensions which are already extensions in the category of finite flat group schemes $\mathbf{Z}_p$. Observe that $\mathrm{Ext}^1_{\mathrm{fl}}(U_{\lambda^n}, U_{\lambda^n}) \cap \mathrm{Ext}^1_{\mathcal{O}[D_p]}(U_{\lambda^n}, U_{\lambda^n})$ is an $\mathcal{O}$-module, so $H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n})$ is seen to be an $\mathcal{O}$-submodule of $H^1(\mathbf{Q}_p, V_{\lambda^n})$. We observe that our definition is equivalent to requiring that the classes in $H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n})$ map under (1.8) to $\mathrm{Ext}^1_{\mathrm{fl}}(U_{\lambda^m}, U_{\lambda^n})$ for all $m \geq n$. For if e_m is the extension class in $\mathrm{Ext}^1(U_{\lambda^m}, U_{\lambda^n})$ then $e_m \hookrightarrow e_n \oplus U_{\lambda^m}$ as Galois-modules and we can apply results of [Ray1] to see that e_m comes from a finite flat group scheme over $\mathbf{Z}_p$ if e_n does.

In the flat (non-ordinary) case $\rho_0|_{I_p}$ is determined by Raynaud's results as mentioned at the beginning of the chapter. It follows in particular that, since $\rho_0|_{D_p}$ is absolutely irreducible, $V(\mathbf{Q}_p) = H^0(\mathbf{Q}_p, V)$ is divisible in this case (in fact $V(\mathbf{Q}_p) \simeq K/\mathcal{O}$). Thus $H^1(\mathbf{Q}_p, V_{\lambda^n}) \simeq H^1(\mathbf{Q}_p, V)_{\lambda^n}$ and hence we can define

$$H^1_{\mathrm{f}}(\mathbf{Q}_p, V) = \bigcup_{n=1}^{\infty} H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n}),$$

and we claim that $H^1_{\mathrm{f}}(\mathbf{Q}_p, V)_{\lambda^n} \simeq H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n})$. To see this we have to compare representations for $m \geq n$,

$$\begin{array}{ccc} \rho_{n,m}\colon \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p) & \longrightarrow & \mathrm{GL}_2(\mathcal{O}_n[\varepsilon]/\lambda^m) \\ \| & & \downarrow \varphi_{m,n} \\ \rho_{m,m}\colon \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p) & \longrightarrow & \mathrm{GL}_2(\mathcal{O}_m[\varepsilon]/\lambda^m) \end{array}$$

where $\rho_{n,m}$ and $\rho_{m,m}$ are obtained from $\alpha_n \in H^1(\mathbf{Q}_p, V_{\lambda^n})$ and $\mathrm{im}(\alpha_n) \in H^1(\mathbf{Q}_p, V_{\lambda^m})$ and $\varphi_{m,n}\colon a + b\varepsilon \to a + \lambda^{m-n}b\varepsilon$. By [Ram, Prop 1.1 and Lemma 2.1] if $\rho_{n,m}$ comes from a finite flat group scheme then so does $\rho_{m,m}$. Conversely $\varphi_{m,n}$ is injective and so $\rho_{n,m}$ comes from a finite flat group scheme if $\rho_{m,m}$ does; cf. [Ray1]. The definitions of $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n})$ and $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$ now extend to the flat case and we note that (1.7) is also valid in the flat case.

Still in the flat (non-ordinary) case we can again use the determination of $\rho_0|_{I_p}$ to see that $H^1(\mathbf{Q}_p, V)$ is divisible. For it is enough to check that $H^2(\mathbf{Q}_p, V_\lambda) = 0$ and this follows by duality from the fact that $H^0(\mathbf{Q}_p, V_\lambda^*) = 0$

where $V_\lambda^* = \text{Hom}(V_\lambda, \boldsymbol{\mu}_p)$ and $\boldsymbol{\mu}_p$ is the group of p^{th} roots of unity. (Again this follows from the explicit form of $\rho_0|_{D_p}$.) Much subtler is the fact that $H^1_{\mathrm{f}}(\mathbf{Q}_p, V)$ is divisible. This result is essentially due to Ramakrishna. For, using a local version of Proposition 1.1 below we have that

$$\text{Hom}_{\mathcal{O}}\,(\mathfrak{p}_R/\mathfrak{p}_R^2,\ K/\mathcal{O}) \simeq H^1_{\mathrm{f}}(\mathbf{Q}_p, V)$$

where R is the universal local flat deformation ring for $\rho_0|_{D_p}$ and $\mathcal{O}$-algebras. (This exists by Theorem 1.1 of [Ram] because $\rho_0|_{D_p}$ is absolutely irreducible.) Since $R \simeq R^{\mathrm{fl}} \underset{W(k)}{\otimes} \mathcal{O}$ where R^{fl} is the corresponding ring for $W(k)$-algebras the main theorem of [Ram, Th. 4.2] shows that R is a power series ring and the divisibility of $H^1_{\mathrm{f}}(\mathbf{Q}_p, V)$ then follows. We refer to [Ram] for more details about R^{fl}.

Next we need an analogue of (1.5) for V. Again this is a variant of standard results in deformation theory and is given (at least for $\mathcal{D} = (\text{ord}, \Sigma, W(k), \phi)$ with some restriction on χ_1, χ_2 in i(a)) in [MT, Prop 25].

PROPOSITION 1.2. *Suppose that $\rho_{f,\lambda}$ is a deformation of ρ_0 of type $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}, \mathcal{M})$ with $\mathcal{O}$ an unramified extension of $\mathcal{O}_{f,\lambda}$. Then as $\mathcal{O}$-modules*

$$\text{Hom}_{\mathcal{O}}(\mathfrak{p}_{\mathcal{D}}/\mathfrak{p}_{\mathcal{D}}^2, K/\mathcal{O}) \simeq H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V).$$

Remark. The isomorphism is functorial in an obvious way if one changes $\mathcal{D}$ to a larger $\mathcal{D}'$.

Proof. We will just describe the Selmer case with $\mathcal{M} = \phi$ as the other cases use similar arguments. Suppose that α is a cocycle which represents a cohomology class in $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n})$. Let $\mathcal{O}_n[\varepsilon]$ denote the ring $\mathcal{O}[\varepsilon]/(\lambda^n\varepsilon, \varepsilon^2)$. We can associate to α a representation

$$\rho_\alpha\colon \text{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \text{GL}_2(\mathcal{O}_n[\varepsilon])$$

as follows: set $\rho_\alpha(g) = \alpha(g)\rho_{f,\lambda}(g)$ where $\rho_{f,\lambda}(g)$, *a priori* in $\text{GL}_2(\mathcal{O})$, is viewed in $\text{GL}_2(\mathcal{O}_n[\varepsilon])$ via the natural mapping $\mathcal{O} \to \mathcal{O}_n[\varepsilon]$. Here a basis for $\mathcal{O}^2$ is chosen so that the representation $\rho_{f,\lambda}$ on the decomposition group $D_p \subset \text{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$ has the upper triangular form of (i)(a), and then $\alpha(g) \in V_{\lambda^n}$ is viewed in $\text{GL}_2(\mathcal{O}_n[\varepsilon])$ by identifying

$$V_{\lambda^n} \simeq \left\{\begin{pmatrix} 1+y\varepsilon & x\varepsilon \\ z\varepsilon & 1-t\varepsilon \end{pmatrix}\right\} = \{\ker : \text{GL}_2\left(\mathcal{O}_n[\varepsilon]\right) \to \text{GL}_2(\mathcal{O})\}.$$

Then

$$W^0_{\lambda^n} = \left\{\begin{pmatrix} 1 & x\varepsilon \\ & 1 \end{pmatrix}\right\},$$

$$W^1_{\lambda^n} = \left\{ \begin{pmatrix} 1+y\varepsilon & x\varepsilon \\ & 1-y\varepsilon \end{pmatrix} \right\},$$

$$W_{\lambda^n} = \left\{ \begin{pmatrix} 1+y\varepsilon & x\varepsilon \\ z\varepsilon & 1-y\varepsilon \end{pmatrix} \right\},$$

and

$$V^1_{\lambda^n} = \left\{ \begin{pmatrix} 1+y\varepsilon & x\varepsilon \\ & 1-t\varepsilon \end{pmatrix} \right\}.$$

One checks readily that ρ_α is a continuous homomorphism and that the deformation $[\rho_\alpha]$ is unchanged if we add a coboundary to α.

We need to check that $[\rho_\alpha]$ is a Selmer deformation. Let $\mathcal{H} = \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p^{\mathrm{unr}})$ and $\mathcal{G} = \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}/\mathbf{Q}_p)$. Consider the exact sequence of $\mathcal{O}[\mathcal{G}]$-modules

$$0 \to (V^1_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{H}} \to (V_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{H}} \to X \to 0$$

where X is a submodule of $(V_{\lambda^n}/V^1_{\lambda^n})^{\mathcal{H}}$. Since the action of D_p on $V_{\lambda^n}/V^1_{\lambda^n}$ is via a character which is nontrivial mod λ (it equals $\chi_2\chi_1^{-1}$ mod λ and $\chi_1 \not\equiv \chi_2$), we see that $X^{\mathcal{G}} = 0$ and $H^1(\mathcal{G}, X) = 0$. Then we have an exact diagram of $\mathcal{O}$-modules

$$\begin{array}{c} 0 \\ \downarrow \\ H^1(\mathcal{G}, (V^1_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{H}}) \simeq H^1(\mathcal{G}, (V_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{H}}) \\ \downarrow \\ H^1(\mathbf{Q}_p, V_{\lambda^n}/W^0_{\lambda^n}) \\ \downarrow \\ H^1(\mathbf{Q}_p^{\mathrm{unr}}, V_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{G}}. \end{array}$$

By hypothesis the image of α is zero in $H^1(\mathbf{Q}_p^{\mathrm{unr}}, V_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{G}}$. Hence it is in the image of $H^1(\mathcal{G}, (V^1_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{H}})$. Thus we can assume that it is represented in $H^1(\mathbf{Q}_p, V_{\lambda^n}/W^0_{\lambda^n})$ by a cocycle, which maps $\mathcal{G}$ to $V^1_{\lambda^n}/W^0_{\lambda^n}$; i.e., $f(D_p) \subset V^1_{\lambda^n}/W^0_{\lambda^n}$, $f(I_p) = 0$. The difference between f and the image of α is a coboundary $\{\sigma \mapsto \sigma\bar{u} - \bar{u}\}$ for some $u \in V_{\lambda^n}$. By subtracting the coboundary $\{\sigma \mapsto \sigma u - u\}$ from α globally we get a new α such that $\alpha = f$ as cocycles mapping $\mathcal{G}$ to $V^1_{\lambda^n}/W^0_{\lambda^n}$. Thus $\alpha(D_p) \subset V^1_{\lambda^n}$, $\alpha(I_p) \subset W^0_{\lambda^n}$ and it is now easy to check that $[\rho_\alpha]$ is a Selmer deformation of ρ_0.

Since $[\rho_\alpha]$ is a Selmer deformation there is a unique map of local $\mathcal{O}$-algebras $\varphi_\alpha : R_{\mathcal{D}} \to \mathcal{O}_n[\varepsilon]$ inducing it. (If $\mathcal{M} \neq \phi$ we must check the

other conditions also.) Since $\rho_\alpha \equiv \rho_{f,\lambda} \bmod \varepsilon$ we see that restricting φ_α to $\mathfrak{p}_\mathcal{D}$ gives a homomorphism of $\mathcal{O}$-modules,

$$\varphi_\alpha : \mathfrak{p}_\mathcal{D} \rightarrow \varepsilon.\mathcal{O}/\lambda^n$$

such that $\varphi_\alpha(\mathfrak{p}_\mathcal{D}^2) = 0$. Thus we have defined a map $\varphi : \alpha \rightarrow \varphi_\alpha$,

$$\varphi : H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n}) \rightarrow \mathrm{Hom}_\mathcal{O}(\mathfrak{p}_\mathcal{D}/\mathfrak{p}_\mathcal{D}^2, \mathcal{O}/\lambda^n).$$

It is straightforward to check that this is a map of $\mathcal{O}$-modules. To check the injectivity of φ suppose that $\varphi_\alpha(\mathfrak{p}_\mathcal{D}) = 0$. Then φ_α factors through $R_\mathcal{D}/\mathfrak{p}_\mathcal{D} \simeq \mathcal{O}$ and being an $\mathcal{O}$-algebra homomorphism this determines φ_α. Thus $[\rho_{f,\lambda}] = [\rho_\alpha]$. If $A^{-1}\rho_\alpha A = \rho_{f,\lambda}$ then $A \bmod \varepsilon$ is seen to be central by Schur's lemma and so may be taken to be I. A simple calculation now shows that α is a coboundary.

To see that φ is surjective choose

$$\Psi \in \mathrm{Hom}_\mathcal{O}(\mathfrak{p}_\mathcal{D}/\mathfrak{p}_\mathcal{D}^2, \mathcal{O}/\lambda^n).$$

Then ρ_Ψ: $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \rightarrow \mathrm{GL}_2(R_\mathcal{D}/(\mathfrak{p}_\mathcal{D}^2, \ker\Psi))$ is induced by a representative of the universal deformation (chosen to equal $\rho_{f,\lambda}$ when reduced mod $\mathfrak{p}_\mathcal{D}$) and we define a map $\alpha_\Psi : \mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \rightarrow V_{\lambda^n}$ by

$$\alpha_\Psi(g) = \rho_\Psi(g)\rho_{f,\lambda}(g)^{-1} \in \left\{ \begin{matrix} 1 + \mathfrak{p}_\mathcal{D}/(\mathfrak{p}_\mathcal{D}^2, \ker\Psi) & \mathfrak{p}_\mathcal{D}/(\mathfrak{p}_\mathcal{D}^2, \ker\Psi) \\ \mathfrak{p}_\mathcal{D}/(\mathfrak{p}_\mathcal{D}^2, \ker\Psi) & 1 + \mathfrak{p}_\mathcal{D}/(\mathfrak{p}_\mathcal{D}^2, \ker\Psi) \end{matrix} \right\} \subseteq V_{\lambda^n}$$

where $\rho_{f,\lambda}(g)$ is viewed in $\mathrm{GL}_2(R_\mathcal{D}/(\mathfrak{p}_\mathcal{D}^2, \ker\Psi))$ via the structural map $\mathcal{O} \rightarrow R_\mathcal{D}$ ($R_\mathcal{D}$ being an $\mathcal{O}$-algebra and the structural map being local because of the existence of a section). The right-hand inclusion comes from

$$\begin{matrix} \mathfrak{p}_D/(\mathfrak{p}_D^2, \ker\Psi) & \overset{\Psi}{\hookrightarrow} & \mathcal{O}/\lambda^n & \overset{\sim}{\rightarrow} & (\mathcal{O}/\lambda^n)\cdot\varepsilon \\ & & 1 & \mapsto & \varepsilon. \end{matrix}$$

Then α_Ψ is readily seen to be a continuous cocycle whose cohomology class lies in $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n})$. Finally $\varphi(\alpha_\Psi) = \Psi$. Moreover, the constructions are compatible with change of n, i.e., for $V_{\lambda^n} \hookrightarrow V_{\lambda^{n+1}}$ and $\lambda : \mathcal{O}/\lambda^n \hookrightarrow \mathcal{O}/\lambda^{n+1}$. □

We now relate the local cohomology groups we have defined to the theory of Fontaine and in particular to the groups of Bloch-Kato [BK]. We will distinguish these by writing H^1_F for the cohomology groups of Bloch-Kato. None of the results described in the rest of this section are used in the rest of the paper. They serve only to relate the Selmer groups we have defined (and later compute) to the more standard versions. Using the lattice associated to $\rho_{f,\lambda}$ we obtain also a lattice $T \simeq \mathcal{O}^4$ with Galois action via $\mathrm{Ad}\,\rho_{f,\lambda}$. Let $\mathcal{V} = T \otimes_{\mathbf{Z}_p} \mathbf{Q}_p$ be the associated vector space and identify V with $\mathcal{V}/T$. Let $\mathrm{pr} : \mathcal{V} \rightarrow V$ be

the natural projection and define cohomology modules by

$$\begin{aligned} H^1_F(\mathbf{Q}_p, \mathcal{V}) &= \ker : H^1(\mathbf{Q}_p, \mathcal{V}) \to H^1(\mathbf{Q}_p, \mathcal{V} \underset{\mathbf{Q}_p}{\otimes} B_{\text{crys}}), \\ H^1_F(\mathbf{Q}_p, V) &= \text{pr}\left(H^1_F(\mathbf{Q}_p, \mathcal{V})\right) \subset H^1(\mathbf{Q}_p, V), \\ H^1_F(\mathbf{Q}_p, V_{\lambda^n}) &= (j_n)^{-1}\left(H^1_F(\mathbf{Q}_p, V)\right) \subset H^1(\mathbf{Q}_p, V_{\lambda^n}), \end{aligned}$$

where $j_n \colon V_{\lambda^n} \to V$ is the natural map and the two groups in the definition of $H^1_F(\mathbf{Q}_p, \mathcal{V})$ are defined using continuous cochains. Similar definitions apply to $\mathcal{V}^* = \text{Hom}_{\mathbf{Q}_p}(\mathcal{V}, \mathbf{Q}_p(1))$ and indeed to any finite-dimensional continuous p-adic representation space. The reader is cautioned that the definition of $H^1_F(\mathbf{Q}_p, V_{\lambda^n})$ is dependent on the lattice T (or equivalently on V). Under certain conditions Bloch and Kato show, using the theory of Fontaine and Lafaille, that this is independent of the lattice (see [BK, Lemmas 4.4 and 4.5]). In any case we will consider in what follows a fixed lattice associated to $\rho = \rho_{f,\lambda}$, $\text{Ad}\,\rho$, etc. Henceforth we will only use the notation $H^1_F(\mathbf{Q}_p, -)$ when the underlying vector space is crystalline.

PROPOSITION 1.3. (i) *If ρ_0 is flat but not ordinary and $\rho_{f,\lambda}$ is associated to a p-divisible group then for all n*

$$H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n}) = H^1_F(\mathbf{Q}_p, V_{\lambda^n}).$$

(ii) *If $\rho_{f,\lambda}$ is ordinary,* $\det \rho_{f,\lambda}\big|_{I_p} = \varepsilon$ *and $\rho_{f,\lambda}$ is associated to a p-divisible group, then for all n,*

$$H^1_F(\mathbf{Q}_p, V_{\lambda^n}) \subseteq H^1_{\mathrm{Se}}(\mathbf{Q}_p, V_{\lambda^n}).$$

Proof. Beginning with (i), we define $H^1_{\mathrm{f}}(\mathbf{Q}_p, \mathcal{V}) = \{\alpha \in H^1(\mathbf{Q}_p, \mathcal{V}) : \kappa(\alpha/\lambda^n) \in H^1_{\mathrm{f}}(\mathbf{Q}_p, V) \text{ for all } n\}$ where $\kappa : H^1(\mathbf{Q}_p, \mathcal{V}) \to H^1(\mathbf{Q}_p, V)$. Then we see that in case (i), $H^1_{\mathrm{f}}(\mathbf{Q}_p, V)$ is divisible. So it is enough to show that

$$H^1_F(\mathbf{Q}_p, \mathcal{V}) = H^1_{\mathrm{f}}(\mathbf{Q}_p, \mathcal{V}).$$

We have to compare two constructions associated to a nonzero element α of $H^1(\mathbf{Q}_p, \mathcal{V})$. The first is to associate an extension

$$0 \to \mathcal{V} \to E \xrightarrow{\delta} K \to 0 \tag{1.9}$$

of K-vector spaces with commuting continuous Galois action. If we fix an e with $\delta(e) = 1$ the action on e is defined by $\sigma e = e + \hat{\alpha}(\sigma)$ with $\hat{\alpha}$ a cocycle representing α. The second construction begins with the image of the subspace $\langle\alpha\rangle$ in $H^1(\mathbf{Q}_p, V)$. By the analogue of Proposition 1.2 in the local case, there is an $\mathcal{O}$-module isomorphism

$$H^1(\mathbf{Q}_p, V) \simeq \text{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, K/\mathcal{O})$$

where R is the universal deformation ring of ρ_0 viewed as a representation of $\mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p)$ on $\mathcal{O}$-algebras and $\mathfrak{p}_R$ is the ideal of R corresponding to $\mathfrak{p}_\mathcal{D}$ (i.e., its inverse image in R). Since $\alpha \neq 0$, associated to $\langle\alpha\rangle$ is a quotient $\mathfrak{p}_R/(\mathfrak{p}_R^2, \mathfrak{a})$ of $\mathfrak{p}_R/\mathfrak{p}_R^2$ which is a free $\mathcal{O}$-module of rank one. We then obtain a homomorphism

$$\rho_\alpha\colon \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p) \rightarrow \mathrm{GL}_2\left(R/(\mathfrak{p}_R^2, \mathfrak{a})\right)$$

induced from the universal deformation (we pick a representation in the universal class). This is associated to an $\mathcal{O}$-module of rank 4 which tensored with K gives a K-vector space $E' \simeq (K)^4$ which is an extension

$$0 \rightarrow \mathcal{U} \rightarrow E' \rightarrow \mathcal{U} \rightarrow 0 \tag{1.10}$$

where $\mathcal{U} \simeq K^2$ has the Galois representation $\rho_{f,\lambda}$ (viewed locally).

In the first construction $\alpha \in H^1_F(\mathbf{Q}_p, \mathcal{V})$ if and only if the extension (1.9) is crystalline, as the extension given in (1.9) is a sum of copies of the more usual extension where $\mathbf{Q}_p$ replaces K in (1.9). On the other hand $\langle\alpha\rangle \subseteq H^1_{\mathrm{f}}(\mathbf{Q}_p, \mathcal{V})$ if and only if the second construction can be made through R^{fl}, or equivalently if and only if E' is the representation associated to a p-divisible group. (*A priori*, the representation associated to ρ_α only has the property that on all finite quotients it comes from a finite flat group scheme. However a theorem of Raynaud [Ray1] says that then ρ_α comes from a p-divisible group. For more details on R^{fl}, the universal flat deformation ring of the local representation ρ_0, see [Ram].) Now the extension E' comes from a p-divisible group if and only if it is crystalline; cf. [Fo, §6]. So we have to show that (1.9) is crystalline if and only if (1.10) is crystalline.

One obtains (1.10) from (1.9) as follows. We view $\mathcal{V}$ as $\mathrm{Hom}_K(\mathcal{U}, \mathcal{U})$ and let

$$X = \ker : \{\mathrm{Hom}_K(\mathcal{U}, \mathcal{U}) \otimes \mathcal{U} \rightarrow \mathcal{U}\}$$

where the map is the natural one $f \otimes w \mapsto f(w)$. (All tensor products in this proof will be as K-vector spaces.) Then as $K[D_p]$-modules

$$E' \simeq (E \otimes \mathcal{U})/X.$$

To check this, one calculates explicitly with the definition of the action on E (given above on e) and on E' (given in the proof of Proposition 1.1). It follows from standard properties of crystalline representations that if E is crystalline, so is $E \otimes \mathcal{U}$ and also E'. Conversely, we can recover E from E' as follows. Consider $E' \otimes \mathcal{U} \simeq (E \otimes \mathcal{U} \otimes \mathcal{U})/(X \otimes \mathcal{U})$. Then there is a natural map $\varphi : E \otimes (\det) \rightarrow E' \otimes \mathcal{U}$ induced by the direct sum decomposition $\mathcal{U} \otimes \mathcal{U} \simeq (\det) \oplus \mathrm{Sym}^2 \mathcal{U}$. Here det denotes a 1-dimensional vector space over K with Galois action via $\det \rho_{f,\lambda}$. Now we claim that φ is injective on $\mathcal{V} \otimes (\det)$. For

if $f \in \mathcal{V}$ then $\varphi(f) = f \otimes (w_1 \otimes w_2 - w_2 \otimes w_1)$ where w_1, w_2 are a basis for $\mathcal{U}$ for which $w_1 \wedge w_2 = 1$ in $\det \simeq K$. So if $\varphi(f) \in X \otimes \mathcal{U}$ then

$$f(w_1) \otimes w_2 - f(w_2) \otimes w_1 = 0 \quad \text{in} \quad \mathcal{U} \otimes \mathcal{U}.$$

But this is false unless $f(w_1) = f(w_2) = 0$ whence $f = 0$. So φ is injective on $\mathcal{V} \otimes \det$ and if φ itself were not injective then E would split contradicting $\alpha \neq 0$. So φ is injective and we have exhibited $E \otimes (\det)$ as a subrepresentation of $E' \otimes \mathcal{U}$ which is crystalline. We deduce that E is crystalline if E' is. This completes the proof of (i).

To prove (ii) we check first that $H^1_{\mathrm{Se}}(\mathbf{Q}_p, V_{\lambda^n}) = j_n^{-1}\left(H^1_{\mathrm{Se}}(\mathbf{Q}_p, V)\right)$ (this was already used in (1.7)). We next have to show that $H^1_F(\mathbf{Q}_p, \mathcal{V}) \subseteq H^1_{\mathrm{Se}}(\mathbf{Q}_p, \mathcal{V})$ where the latter is defined by

$$H^1_{\mathrm{Se}}(\mathbf{Q}_p, \mathcal{V}) = \ker : H^1(\mathbf{Q}_p, \mathcal{V}) \to H^1(\mathbf{Q}_p^{\mathrm{unr}}, \mathcal{V}/\mathcal{V}^0)$$

with $\mathcal{V}^0$ the subspace of $\mathcal{V}$ on which I_p acts via ε. But this follows from the computations in Corollary 3.8.4 of [BK]. Finally we observe that

$$\mathrm{pr}\left(H^1_{\mathrm{Se}}(\mathbf{Q}_p, \mathcal{V})\right) \subseteq H^1_{\mathrm{Se}}(\mathbf{Q}_p, V)$$

although the inclusion may be strict, and

$$\mathrm{pr}\left(H^1_F(\mathbf{Q}_p, \mathcal{V})\right) = H^1_F(\mathbf{Q}_p, V)$$

by definition. This completes the proof. □

These groups have the property that for $s \geq r$,

$$H^1(\mathbf{Q}_p, V_{\lambda^r}) \cap j_{r,s}^{-1}\left(H^1_F(\mathbf{Q}_p, V_{\lambda^s})\right) = H^1_F(\mathbf{Q}_p, V_{\lambda^r}) \tag{1.11}$$

where $j_{r,s} : V_{\lambda^r} \to V_{\lambda^s}$ is the natural injection. The same holds for $V^*_{\lambda^r}$ and $V^*_{\lambda^s}$ in place of V_{λ^r} and V_{λ^s} where $V^*_{\lambda^r}$ is defined by

$$V^*_{\lambda^r} = \mathrm{Hom}(V_{\lambda^r}, \boldsymbol{\mu}_{p^r})$$

and similarly for $V^*_{\lambda^s}$. Both results are immediate from the definition (and indeed were part of the motivation for the definition).

We also give a finite level version of a result of Bloch-Kato which is easily deduced from the vector space version. As before let $T \subset \mathcal{V}$ be a Galois stable lattice so that $T \simeq \mathcal{O}^4$. Define

$$H^1_F(\mathbf{Q}_p, T) = i^{-1}\left(H^1_F(\mathbf{Q}_p, \mathcal{V})\right)$$

under the natural inclusion $i : T \hookrightarrow \mathcal{V}$, and likewise for the dual lattice $T^* = \mathrm{Hom}_{\mathbf{Z}_p}(V, (\mathbf{Q}_p/\mathbf{Z}_p)(1))$ in $\mathcal{V}^*$. (Here $\mathcal{V}^* = \mathrm{Hom}(\mathcal{V}, \mathbf{Q}_p(1))$; throughout this paper we use M^* to denote a dual of M with a Cartier twist.) Also write

$\mathrm{pr}_n : T \to T/\lambda^n$ for the natural projection map, and for the mapping it induces on cohomology.

PROPOSITION 1.4. *If $\rho_{f,\lambda}$ is associated to a p-divisible group (the ordinary case is allowed) then*

(i) $\mathrm{pr}_n\left(H^1_F(\mathbf{Q}_p, T)\right) = H^1_F(\mathbf{Q}_p, T/\lambda^n)$ *and similarly for* $T^*, T^*/\lambda^n$.

(ii) $H^1_F(\mathbf{Q}_p, V_{\lambda^n})$ *is the orthogonal complement of* $H^1_F(\mathbf{Q}_p, V^*_{\lambda^n})$ *under Tate local duality between* $H^1(\mathbf{Q}_p, V_{\lambda^n})$ *and* $H^1(\mathbf{Q}_p, V^*_{\lambda^n})$ *and similarly for* W_{λ^n} *and* $W^*_{\lambda^n}$ *replacing* V_{λ^n} *and* $V^*_{\lambda^n}$.

More generally these results hold for any crystalline representation $\mathcal{V}'$ in place of $\mathcal{V}$ and λ' a uniformizer in K' where K' is any finite extension of $\mathbf{Q}_p$ with $K' \subset \mathrm{End}_{\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)} \mathcal{V}'$.

Proof. We first observe that $\mathrm{pr}_n(H^1_F(\mathbf{Q}_p, T)) \subset H^1_F(\mathbf{Q}_p, T/\lambda^n)$. Now from the construction we may identify T/λ^n with V_{λ^n}. A result of Bloch-Kato ([BK, Prop. 3.8]) says that $H^1_F(\mathbf{Q}_p, \mathcal{V})$ and $H^1_F(\mathbf{Q}_p, \mathcal{V}^*)$ are orthogonal complements under Tate local duality. It follows formally that $H^1_F(\mathbf{Q}_p, V^*_{\lambda^n})$ and $\mathrm{pr}_n(H^1_F(\mathbf{Q}_p, T))$ are orthogonal complements, so to prove the proposition it is enough to show that

$$\# H^1_F(\mathbf{Q}_p, V^*_{\lambda^n}) \# H^1_F(\mathbf{Q}_p, V_{\lambda^n}) = \# H^1(\mathbf{Q}_p, V_{\lambda^n}). \tag{1.12}$$

Now if $r = \dim_K H^1_F(\mathbf{Q}_p, \mathcal{V})$ and $s = \dim_K H^1_F(\mathbf{Q}_p, \mathcal{V}^*)$ then

$$r + s = \dim_K H^0(\mathbf{Q}_p, \mathcal{V}) + \dim_K H^0(\mathbf{Q}_p, \mathcal{V}^*) + \dim_K \mathcal{V}. \tag{1.13}$$

From the definition,

$$\# H^1_F(\mathbf{Q}_p, V_{\lambda^n}) = \# (\mathcal{O}/\lambda^n)^r \cdot \# \ker\{H^1(\mathbf{Q}_p, V_{\lambda^n}) \to H^1(\mathbf{Q}_p, V)\}. \tag{1.14}$$

The second factor is equal to $\# \{V(\mathbf{Q}_p)/\lambda^n V(\mathbf{Q}_p)\}$. When we write $V(\mathbf{Q}_p)^{\mathrm{div}}$ for the maximal divisible subgroup of $V(\mathbf{Q}_p)$ this is the same as

$$\begin{aligned} \# (V(\mathbf{Q}_p)/V(\mathbf{Q}_p)^{\mathrm{div}})/\lambda^n &= \# (V(\mathbf{Q}_p)/V(\mathbf{Q}_p)^{\mathrm{div}})_{\lambda^n} \\ &= \# V(\mathbf{Q}_p)_{\lambda^n} / \# (V(\mathbf{Q}_p)^{\mathrm{div}})_{\lambda^n}. \end{aligned}$$

Combining this with (1.14) gives

$$\begin{aligned} \# H^1_F(\mathbf{Q}_p, V_{\lambda^n}) &= \# (\mathcal{O}/\lambda^n)^r \\ &\quad \cdot \# H^0(\mathbf{Q}_p, V_{\lambda^n}) \,/\, \# (\mathcal{O}/\lambda^n)^{\dim_K H^0(\mathbf{Q}_p, \mathcal{V})}. \end{aligned} \tag{1.15}$$

This, together with an analogous formula for $\# H^1_F(\mathbf{Q}_p, V^*_{\lambda^n})$ and (1.13), gives

$$\# H^1_F(\mathbf{Q}_p, V_{\lambda^n}) \# H^1_F(\mathbf{Q}_p, V^*_{\lambda^n}) = \# (\mathcal{O}/\lambda^n)^4 \cdot \# H^0(\mathbf{Q}_p, V_{\lambda^n}) \# H^0(\mathbf{Q}_p, V^*_{\lambda^n}).$$

As $\# H^0(\mathbf{Q}_p, V^*_{\lambda^n}) = \# H^2(\mathbf{Q}_p, V_{\lambda^n})$ the assertion of (1.12) now follows from the formula for the Euler characteristic of V_{λ^n}.

The proof for W_{λ^n}, or indeed more generally for any crystalline representation, is the same. □

We also give a characterization of the orthogonal complements of $H^1_{\mathrm{Se}}(\mathbf{Q}_p, W_{\lambda^n})$ and $H^1_{\mathrm{Se}}(\mathbf{Q}_p, V_{\lambda^n})$, under Tate's local duality. We write these duals as $H^1_{\mathrm{Se}^*}(\mathbf{Q}_p, W^*_{\lambda^n})$ and $H^1_{\mathrm{Se}^*}(\mathbf{Q}_p, V^*_{\lambda^n})$ respectively. Let

$$\varphi_w : H^1(\mathbf{Q}_p, W^*_{\lambda^n}) \to H^1(\mathbf{Q}_p, W^*_{\lambda^n}/(W^*_{\lambda^n})^0)$$

be the natural map where $(W^*_{\lambda^n})^i$ is the orthogonal complement of $W^{1-i}_{\lambda^n}$ in $W^*_{\lambda^n}$, and let $X_{n,i}$ be defined as the image under the composite map

$$\begin{aligned} X_{n,i} = \mathrm{im} : \mathbf{Z}_p^\times/(\mathbf{Z}_p^\times)^{p^n} \otimes \mathcal{O}/\lambda^n &\to H^1(\mathbf{Q}_p, \boldsymbol{\mu}_{p^n} \otimes \mathcal{O}/\lambda^n) \\ &\to H^1(\mathbf{Q}_p, W^*_{\lambda^n}/(W^*_{\lambda^n})^0) \end{aligned}$$

where in the middle term $\boldsymbol{\mu}_{p^n} \otimes \mathcal{O}/\lambda^n$ is to be identified with $(W^*_{\lambda^n})^1/(W^*_{\lambda^n})^0$. Similarly if we replace $W^*_{\lambda^n}$ by $V^*_{\lambda^n}$ we let $Y_{n,i}$ be the image of $\mathbf{Z}_p^\times/(\mathbf{Z}_p^\times)^{p^n} \otimes (\mathcal{O}/\lambda^n)^2$ in $H^1(\mathbf{Q}_p, V^*_{\lambda^n}/(W^*_{\lambda^n})^0)$, and we replace φ_w by the analogous map φ_v.

PROPOSITION 1.5.

$$\begin{aligned} H^1_{\mathrm{Se}^*}(\mathbf{Q}_p, W^*_{\lambda^n}) &= \varphi_w^{-1}(X_{n,i}), \\ H^1_{\mathrm{Se}^*}(\mathbf{Q}_p, V^*_{\lambda^n}) &= \varphi_v^{-1}(Y_{n,i}). \end{aligned}$$

Proof. This can be checked by dualizing the sequence

$$\begin{aligned} 0 &\to H^1_{\mathrm{Str}}(\mathbf{Q}_p, W_{\lambda^n}) \to H^1_{\mathrm{Se}}(\mathbf{Q}_p, W_{\lambda^n}) \\ &\to \ker : \{H^1(\mathbf{Q}_p, W_{\lambda^n}/(W_{\lambda^n})^0) \to H^1(\mathbf{Q}_p^{\mathrm{unr}}, W_{\lambda^n}/(W_{\lambda^n})^0\}, \end{aligned}$$

where $H^1_{\mathrm{str}}(\mathbf{Q}_p, W_{\lambda^n}) = \ker : H^1(\mathbf{Q}_p, W_{\lambda^n}) \to H^1(\mathbf{Q}_p, W_{\lambda^n}/(W_{\lambda^n})^0)$. The first term is orthogonal to $\ker : H^1(\mathbf{Q}_p, W^*_{\lambda^n}) \to H^1(\mathbf{Q}_p, W^*_{\lambda^n}/(W^*_{\lambda^n})^1)$. By the naturality of the cup product pairing with respect to quotients and subgroups the claim then reduces to the well known fact that under the cup product pairing

$$H^1(\mathbf{Q}_p, \boldsymbol{\mu}_{p^n}) \times H^1(\mathbf{Q}_p, \mathbf{Z}/p^n) \to \mathbf{Z}/p^n$$

the orthogonal complement of the unramified homomorphisms is the image of the units $\mathbf{Z}_p^\times/(\mathbf{Z}_p^\times)^{p^n} \to H^1(\mathbf{Q}_p, \boldsymbol{\mu}_{p^n})$. The proof for V_{λ^n} is essentially the same. □

2. Some computations of cohomology groups

We now make some comparisons of orders of cohomology groups using the theorems of Poitou and Tate. We retain the notation and conventions of Section 1 though it will be convenient to state the first two propositions in a more general context. Suppose that

$$L = \prod L_q \subseteq \prod_{q\in\Sigma} H^1(\mathbf{Q}_q, X)$$

is a subgroup, where X is a finite module for $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$ of p-power order. We define L^* to be the orthogonal complement of L under the perfect pairing (local Tate duality)

$$\prod_{q\in\Sigma} H^1(\mathbf{Q}_q, X) \times \prod_{q\in\Sigma} H^1(\mathbf{Q}_q, X^*) \rightarrow \mathbf{Q}_p/\mathbf{Z}_p$$

where $X^* = \mathrm{Hom}(X, \boldsymbol{\mu}_{p^\infty})$. Let

$$\lambda_X : H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, X) \rightarrow \prod_{q\in\Sigma} H^1(\mathbf{Q}_q, X)$$

be the localization map and similarly λ_{X^*} for X^*. Then we set

$$H^1_L(\mathbf{Q}_\Sigma/\mathbf{Q}, X) = \lambda_X^{-1}(L), \quad H^1_{L^*}(\mathbf{Q}_\Sigma/\mathbf{Q}, X^*) = \lambda_{X^*}^{-1}(L^*).$$

The following result was suggested by a result of Greenberg (cf. [Gre1]) and is a simple consequence of the theorems of Poitou and Tate. Recall that p is always assumed odd and that $p \in \Sigma$.

PROPOSITION 1.6.

$$\#H^1_L(\mathbf{Q}_\Sigma/\mathbf{Q},X) \,/\, \#H^1_{L^*}(\mathbf{Q}_\Sigma/\mathbf{Q},X^*) = h_\infty \prod_{q\in\Sigma} h_q$$

where

$$\begin{cases} h_q = \#H^0(\mathbf{Q}_q,X^*)/[H^1(\mathbf{Q}_q,X):L_q] \\ h_\infty = \#H^0(\mathbf{R}, X^*)\,\#H^0(\mathbf{Q},X)/\#H^0(\mathbf{Q},X^*). \end{cases}$$

Proof. Adapting the exact sequence of Poitou and Tate (cf. [Mi2, Th. 4.20]) we get a seven term exact sequence

$$\begin{array}{ccccccc} 0 & \longrightarrow & H^1_L(\mathbf{Q}_\Sigma/\mathbf{Q}, X) & \longrightarrow & H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, X) & \longrightarrow & \prod_{q\in\Sigma} H^1(\mathbf{Q}_q, X)/L_q \\ & & & & & & \downarrow \\ \prod_{q\in\Sigma} H^2(\mathbf{Q}_q, X) & \longleftarrow & H^2(\mathbf{Q}_\Sigma/\mathbf{Q}, X) & \longleftarrow & H^1_{L^*}(\mathbf{Q}_\Sigma/\mathbf{Q}, X^*)^\wedge & & \end{array}$$

$$\hookrightarrow H^0(\mathbf{Q}_\Sigma/\mathbf{Q}, X^*)^\wedge \longrightarrow 0,$$

where $M^\wedge = \mathrm{Hom}(M, \mathbf{Q}_p/\mathbf{Z}_p)$. Now using local duality and global Euler characteristics (cf. [Mi2, Cor. 2.3 and Th. 5.1]) we easily obtain the formula in the proposition. We repeat that in the above proposition X can be arbitrary of p-power order. $\square$

We wish to apply the proposition to investigate $H^1_{\mathcal{D}}$. Let $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}, \mathcal{M})$ be a standard deformation theory as in Section 1 and define a corresponding group $L_n = L_{\mathcal{D},n}$ by setting

$$L_{n,q} = \begin{cases} H^1(\mathbf{Q}_q, V_{\lambda^n}) & \text{for } q \neq p \quad \text{and} \quad q \notin \mathcal{M} \\ H^1_{D_q}(\mathbf{Q}_q, V_{\lambda^n}) & \text{for } q \neq p \quad \text{and} \quad q \in \mathcal{M} \\ H^{.1}(\mathbf{Q}_p, V_{\lambda^n}) & \text{for } q = p. \end{cases}$$

Then $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n}) = H^1_{L_n}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_{\lambda^n})$ and we also define

$$H^1_{\mathcal{D}^*}(\mathbf{Q}_\Sigma/\mathbf{Q}, V^*_{\lambda^n}) = H^1_{L^*_n}(\mathbf{Q}_\Sigma/\mathbf{Q}, V^*_{\lambda^n}).$$

We will adopt the convention implicit in the above that if we consider $\Sigma' \supset \Sigma$ then $H^1_{\mathcal{D}}(\mathbf{Q}_{\Sigma'}/\mathbf{Q}, V_{\lambda^n})$ places no local restriction on the cohomology classes at primes $q \in \Sigma' - \Sigma$. Thus in $H^1_{\mathcal{D}^*}(\mathbf{Q}_{\Sigma'}/\mathbf{Q}, V^*_{\lambda^n})$ we will require (by duality) that the cohomology class be locally trivial at $q \in \Sigma' - \Sigma$.

We need now some estimates for the local cohomology groups. First we consider an arbitrary finite $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$-module X:

PROPOSITION 1.7. *If $q \notin \Sigma$, and X is an arbitrary finite* $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$-*module of p-power order,*

$$\#H^1_{L'}(\mathbf{Q}_{\Sigma\cup q}/\mathbf{Q}, X)/\#H^1_L(\mathbf{Q}_\Sigma/\mathbf{Q}, X) \leq \#H^0(\mathbf{Q}_q, X^*)$$

where $L'_\ell = L_\ell$ for $\ell \in \Sigma$ and $L'_q = H^1(\mathbf{Q}_q, X)$.

Proof. Consider the short exact sequence of inflation-restriction:

$$\begin{array}{ccccc} 0 \to H^1_L(\mathbf{Q}_\Sigma/\mathbf{Q}, X) \to & H^1_{L'}(\mathbf{Q}_{\Sigma\cup q}/\mathbf{Q}, X) & \to & \mathrm{Hom}(\mathrm{Gal}(\mathbf{Q}_{\Sigma\cup q}/\mathbf{Q}_\Sigma), X)^{\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})} \\ & \Big\downarrow & & \cap\Big\downarrow \\ & H^1(\mathbf{Q}_q^{\mathrm{unr}}, X)^{\mathrm{Gal}(\mathbf{Q}_q^{\mathrm{unr}}/\mathbf{Q}_q)} & \xrightarrow{\sim} & H^1(\mathbf{Q}_q^{\mathrm{unr}}, X)^{\mathrm{Gal}(\mathbf{Q}_q^{\mathrm{unr}}/\mathbf{Q}_q)} \end{array}$$

The proposition follows when we note that

$$\#H^0(\mathbf{Q}_q, X^*) \quad = \quad \#H^1(\mathbf{Q}_q^{\mathrm{unr}}, X)^{\mathrm{Gal}(\mathbf{Q}_q^{\mathrm{unr}}/\mathbf{Q}_q)}. \qquad \square$$

Now we return to the study of V_{λ^n} and W_{λ^n}.

PROPOSITION 1.8. *If $q \in \mathcal{M}$ $(q \neq p)$ and $X = V_{\lambda^n}$ then $h_q = 1$.*

Proof. This is a straightforward calculation. For example if q is of type (A) then we have

$$L_{n,q} = \ker\{H^1(\mathbf{Q}_q, V_{\lambda^n}) \;\to\; H^1(\mathbf{Q}_q, W_{\lambda^n}/W^0_{\lambda^n}) \oplus H^1(\mathbf{Q}_q^{\mathrm{unr}},\; \mathcal{O}/\lambda^n)\}.$$

Using the long exact sequence of cohomology associated to

$$0 \to W^0_{\lambda^n} \to W_{\lambda^n} \to W_{\lambda^n}/W^0_{\lambda^n} \to 0$$

one obtains a formula for the order of $L_{n,q}$ in terms of $\#H^i(\mathbf{Q}_q, W_{\lambda^n})$, $\#H^i(\mathbf{Q}_q, W_{\lambda^n}/W^0_{\lambda^n})$ etc. Using local Euler characteristics these are easily reduced to ones involving $H^0(\mathbf{Q}_q, W^*_{\lambda^n})$ etc. and the result follows easily. □

The calculation of h_p is more delicate. We content ourselves with an inequality in some cases.

PROPOSITION 1.9. (i) *If* $X = V_{\lambda^n}$ *then*

$$h_p\, h_\infty = \#(\mathcal{O}/\lambda)^{3n}\, \# H^0(\mathbf{Q}_p, V^*_{\lambda^n}) / \# H^0(\mathbf{Q}, V^*_{\lambda^n})$$

in the unrestricted case.

(ii) *If* $X = V_{\lambda^n}$ *then*

$$h_p\, h_\infty \le \#(\mathcal{O}/\lambda)^{n}\, \# H^0(\mathbf{Q}_p, (V^{\mathrm{ord}}_{\lambda^n})^*) / \# H^0(\mathbf{Q}, W^*_{\lambda^n})$$

in the ordinary case.

(iii) *If* $X = V_{\lambda^n}$ *or* W_{λ^n} *then* $h_p\, h_\infty \le \# H^0(\mathbf{Q}_p, (W^0_{\lambda^n})^*) / \# H^0(\mathbf{Q}, W^*_{\lambda^n})$ *in the Selmer case.*

(iv) *If* $X = V_{\lambda^n}$ *or* W_{λ^n} *then* $h_p\, h_\infty = 1$ *in the strict case.*

(v) *If* $X = V_{\lambda^n}$ *then* $h_p\, h_\infty = 1$ *in the flat case.*

(vi) *If* $X = V_{\lambda^n}$ *or* W_{λ^n} *then* $h_p\, h_\infty = 1/\# H^0(\mathbf{Q}, V^*_{\lambda^n})$ *if* $L_{n,p} = H^1_F(\mathbf{Q}_p, X)$ *and* $\rho_{f,\lambda}$ *arises from an ordinary* p*-divisible group.*

Proof. Case (i) is trivial. Consider then case (iii) with $X = V_{\lambda^n}$. We have a long exact sequence of cohomology associated to the exact sequence:

$$0 \to W^0_{\lambda^n} \to V_{\lambda^n} \to V_{\lambda^n}/W^0_{\lambda^n} \to 0. \tag{1.16}$$

In particular this gives the map u in the diagram

$$\begin{array}{ccc} & H^1(\mathbf{Q}_p, V_{\lambda^n}) & \\ & \downarrow u & \searrow \delta \\ 1 \to Z = H^1(\mathbf{Q}_p^{\mathrm{unr}}/\mathbf{Q}_p, (V_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{H}}) \to & H^1(\mathbf{Q}_p, V_{\lambda^n}/W^0_{\lambda^n}) & \to H^1(\mathbf{Q}_p^{\mathrm{unr}}, V_{\lambda^n}/W^0_{\lambda^n})^{\mathcal{G}} \to 1 \end{array}$$

where $\mathcal{G} = \mathrm{Gal}(\mathbf{Q}_p^{\mathrm{unr}}/\mathbf{Q}_p)$, $\mathcal{H} = \mathrm{Gal}(\bar{\mathbf{Q}}_p/\mathbf{Q}_p^{\mathrm{unr}})$ and δ is defined to make the triangle commute. Then writing $h_i(M)$ for $\#H^i(\mathbf{Q}_p, M)$ we have that $\#Z =$

$h_0(V_{\lambda^n}/W^0_{\lambda^n})$ and $\#\operatorname{im}\delta \geq (\#\operatorname{im}u)/(\#Z)$. A simple calculation using the long exact sequence associated to (1.16) gives

$$\#\operatorname{im}u = \frac{h_1(V_{\lambda^n}/W^0_{\lambda^n})h_2(V_{\lambda^n})}{h_2(W^0_{\lambda^n})h_2(V_{\lambda^n}/W^0_{\lambda^n})}\,. \tag{1.17}$$

Hence

$$[H^1(\mathbf{Q}_p, V_{\lambda^n}) \colon L_{n,p}] = \#\operatorname{im}\delta \geq \#(\mathcal{O}/\lambda)^{3n} h_0(V^*_{\lambda^n})/h_0(W^0_{\lambda^n}{}^*).$$

The inequality in (iii) follows for $X = V_{\lambda^n}$ and the case $X = W_{\lambda^n}$ is similar. Case (ii) is similar. In case (iv) we just need $\#\operatorname{im}u$ which is given by (1.17) with W_{λ^n} replacing V_{λ^n}. In case (v) we have already observed in Section 1 that Raynaud's results imply that $\# H^0(\mathbf{Q}_p, V^*_{\lambda^n}) = 1$ in the flat case. Moreover $\# H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n})$ can be computed to be $\#(\mathcal{O}/\lambda)^{2n}$ from

$$H^1_{\mathrm{f}}(\mathbf{Q}_p, V_{\lambda^n}) \simeq H^1_{\mathrm{f}}(\mathbf{Q}_p, V)_{\lambda^n} \simeq \operatorname{Hom}_{\mathcal{O}}(\mathfrak{p}_R/\mathfrak{p}_R^2, K/\mathcal{O})_{\lambda^n}$$

where R is the universal local flat deformation ring of ρ_0 for $\mathcal{O}$-algebras. Using the relation $R \simeq R^{\mathrm{fl}} \underset{W(k)}{\otimes} \mathcal{O}$ where R^{fl} is the corresponding ring for $W(k)$-algebras, and the main theorem of [Ram] (Theorem 4.2) which computes R^{fl}, we can deduce the result.

We now prove (vi). From the definitions

$$\# H^1_F(\mathbf{Q}_p, V_{\lambda^n}) = \begin{cases} (\#\mathcal{O}/\lambda^n)^r \ \#H^0(\mathbf{Q}_p, W_{\lambda^n}) & \text{if } \rho_{f,\lambda}|_{D_p} \text{ does not split} \\ (\#\mathcal{O}/\lambda^n)^r & \text{if } \rho_{f,\lambda}|_{D_p} \text{ splits} \end{cases}$$

where $r = \dim_K H^1_F(\mathbf{Q}_p, \mathcal{V})$. This we can compute using the calculations in [BK, Cor. 3.8.4]. We find that $r = 2$ in the non-split case and $r = 3$ in the split case and (vi) follows easily. □

3. Some results on subgroups of $\mathrm{GL}_2(k)$

We now give two group-theoretic results which will not be used until Chapter 3. Although these could be phrased in purely group-theoretic terms it will be more convenient to continue to work in the setting of Section 1, i.e., with ρ_0 as in (1.1) so that $\operatorname{im}\rho_0$ is a subgroup of $\mathrm{GL}_2(k)$ and $\det\rho_0$ is assumed odd.

LEMMA 1.10. *If* $\operatorname{im}\rho_0$ *has order divisible by* p *then:*

(i) *It contains an element* γ_0 *of order* $m \geq 3$ *with* $(m,p) = 1$ *and* γ_0 *trivial on any abelian quotient of* $\operatorname{im}\rho_0$.

(ii) *It contains an element* $\rho_0(\sigma)$ *with any prescribed image in the Sylow 2-subgroup of* $(\operatorname{im}\rho_0)/(\operatorname{im}\rho_0)'$ *and with the ratio of the eigenvalues not equal to* $\omega(\sigma)$. *(Here* $(\operatorname{im}\rho_0)'$ *denotes the derived subgroup of* $(\operatorname{im}\rho_0)$.*)*

The same results hold if the image of the projective representation $\tilde{\rho}_0$ associated to ρ_0 is isomorphic to A_4, S_4 or A_5.

Proof. (i) Let $G = \operatorname{im}\rho_0$ and let Z denote the center of G. Then we have a surjection $G' \longrightarrow (G/Z)'$ where the $'$ denotes the derived group. By Dickson's classification of the subgroups of $\mathrm{GL}_2(k)$ containing an element of order p, (G/Z) is isomorphic to $\mathrm{PGL}_2(k')$ or $\mathrm{PSL}_2(k')$ for some finite field k' of characteristic p or possibly to A_5 when $p = 3$, cf. [Di, §260]. In each case we can find, and then lift to G', an element of order m with $(m, p) = 1$ and $m \geq 3$, except possibly in the case $p = 3$ and $\mathrm{PSL}_2(\mathbf{F}_3) \simeq A_4$ or $\mathrm{PGL}_2(\mathbf{F}_3) \simeq S_4$. However in these cases $(G/Z)'$ has order divisible by 4 so the 2-Sylow subgroup of G' has order greater than 2. Since it has at most one element of exact order 2 (the eigenvalues would both be -1 since it is in the kernel of the determinant and hence the element would be $-I$) it must also have an element of order 4.

The argument in the A_4, S_4 and A_5 cases is similar.

(ii) Since ρ_0 is assumed absolutely irreducible, $G = \operatorname{im}\rho_0$ has no fixed line. We claim that the same then holds for the derived group G'. For otherwise since $G' \triangleleft G$ we could obtain a second fixed line by taking $\langle gv \rangle$ where $\langle v \rangle$ is the original fixed line and g is a suitable element of G. Thus G' would be contained in the group of diagonal matrices for a suitable basis and either it would be central in which case G would be abelian or its normalizer in $\mathrm{GL}_2(k)$, and hence also G, would have order prime to p. Since neither of these possibilities is allowed, G' has no fixed line.

By Dickson's classification of the subgroups of $\mathrm{GL}_2(k)$ containing an element of order p the image of $\operatorname{im}\rho_0$ in $\mathrm{PGL}_2(k)$ is isomorphic to $\mathrm{PGL}_2(k')$ or $\mathrm{PSL}_2(k')$ for some finite field k' of characteristic p or possibly to A_5 when $p = 3$. The only one of these with a quotient group of order p is $\mathrm{PSL}_2(\mathbf{F}_3)$ when $p = 3$. It follows that $p \nmid [G : G']$ except in this one case which we treat separately. So assuming now that $p \nmid [G : G']$ we see that G' contains a nontrivial unipotent element u. Since G' has no fixed line there must be another noncommuting unipotent element v in G'. Pick a basis for $\rho_0|_{G'}$ consisting of their fixed vectors. Then let τ be an element of $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q})$ for which the image of $\rho_0(\tau)$ in G/G' is prescribed and let $\rho_0(\tau) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Then

$$\delta = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & s\alpha \\ & 1 \end{pmatrix} \begin{pmatrix} 1 & \\ r\beta & 1 \end{pmatrix}$$

has $\det(\delta) = \det \rho_0(\tau)$ and $\operatorname{trace}\delta = s\alpha(ra\beta + c) \ + br\beta + a + d$. Since $p \geq 3$ we can choose this trace to avoid any two given values (by varying s) unless $ra\beta + c = 0$ for all r. But $ra\beta + c$ cannot be zero for all r as otherwise $a = c = 0$. So we can find a δ for which the ratio of the eigenvalues is not $\omega(\tau)$, $\det(\delta)$ being, of course, fixed.

Now suppose that $\operatorname{im}\rho_0$ does not have order divisible by p but that the associated projective representation $\widetilde{\rho_0}$ has image isomorphic to S_4 or A_5, so necessarily $p \neq 3$. Pick an element τ such that the image of $\rho_0(\tau)$ in G/G' is any prescribed class. Since this fixes both $\det\rho_0(\tau)$ and $\omega(\tau)$ we have to show that we can avoid at most two particular values of the trace for τ. To achieve this we can adapt our first choice of τ by multiplying by any element of G'. So pick $\sigma \in G'$ as in (i) which we can assume in these two cases has order 3. Pick a basis for ρ_0, by extending scalars if necessary, so that $\sigma \mapsto \left(\begin{smallmatrix}\alpha & \\ & \alpha^{-1}\end{smallmatrix}\right)$. Then one checks easily that if $\rho_0(\tau) = \left(\begin{smallmatrix}a & b\\ c & d\end{smallmatrix}\right)$ we cannot have the traces of all of τ, $\sigma\tau$ and $\sigma^2\tau$ lying in a set of the form $\{\mp t\}$ unless $a = d = 0$. However we can ensure that $\rho_0(\tau)$ does not satisfy this by first multiplying τ by a suitable element of G' since G' is not contained in the diagonal matrices (it is not abelian).

In the A_4 case, and in the $\mathrm{PSL}_2(\mathbf{F}_3) \simeq A_4$ case when $p = 3$, we use a different argument. In both cases we find that the 2-Sylow subgroup of G/G' is generated by an element z in the centre of G. Either a power of z is a suitable candidate for $\rho_0(\sigma)$ or else we must multiply the power of z by an element of G', the ratio of whose eigenvalues is not equal to 1. Such an element exists because in G' the only possible elements without this property are $\{\mp I\}$ (such elements necessarily have determinant 1 and order prime to p) and we know that $\#G' > 2$ as was noted in the proof of part (i). □

Remark. By a well-known result on the finite subgroups of $\mathrm{PGL}_2(\overline{\mathbf{F}}_p)$ this lemma covers all ρ_0 whose images are absolutely irreducible and for which $\widetilde{\rho_0}$ is not dihedral.

Let K_1 be the splitting field of ρ_0. Then we can view W_λ and W_λ^* as $\mathrm{Gal}(K_1(\zeta_p)/\mathbf{Q})$-modules. We need to analyze their cohomology. Recall that we are assuming that ρ_0 is absolutely irreducible. Let $\widetilde{\rho_0}$ be the associated projective representation to $\mathrm{PGL}_2(k)$.

The following proposition is based on the computations in [CPS].

PROPOSITION 1.11. *Suppose that ρ_0 is absolutely irreducible. Then*

$$H^1(K_1(\zeta_p)/\mathbf{Q}, W_\lambda^*) = 0.$$

Proof. If the image of ρ_0 has order prime to p the lemma is trivial. The subgroups of $\mathrm{GL}_2(k)$ containing an element of order p which are not contained in a Borel subgroup have been classified by Dickson [Di, §260] or [Hu, II.8.27] Their images inside $\mathrm{PGL}_2(k')$ where k' is the quadratic extension of k are conjugate to $\mathrm{PGL}_2(F)$ or $\mathrm{PSL}_2(F)$ for some subfield F of k', or they are isomorphic to one of the exceptional groups A_4, S_4, A_5.

Assume then that the cohomology group $H^1(K_1(\zeta_p)/\mathbf{Q},\, W_\lambda^*) \neq 0$. Then by considering the inflation-restriction sequence with respect to the normal

subgroup $\mathrm{Gal}(K_1(\zeta_p)/K_1)$ we see that $\zeta_p \in K_1$. Next, since the representation is (absolutely) irreducible, the center Z of $\mathrm{Gal}(K_1/\mathbf{Q})$ is contained in the diagonal matrices and so acts trivially on W_λ. So by considering the inflation-restriction sequence with respect to Z we see that Z acts trivially on ζ_p (and on W_λ^*). So $\mathrm{Gal}(\mathbf{Q}(\zeta_p)/\mathbf{Q})$ is a quotient of $\mathrm{Gal}(K_1/\mathbf{Q})/Z$. This rules out all cases when $p \neq 3$, and when $p = 3$ we only have to consider the case where the image of the projective representation is isomorphic as a group to $\mathrm{PGL}_2(F)$ for some finite field of characteristic 3. (Note that $S_4 \simeq \mathrm{PGL}_2(\mathbf{F}_3)$.)

Extending scalars commutes with formation of duals and H^1, so we may assume without loss of generality $F \subseteq k$. If $p = 3$ and $\#F > 3$ then $H^1(\mathrm{PSL}_2(F), W_\lambda) = 0$ by results of [CPS]. Then if $\widetilde{\rho_0}$ is the projective representation associated to ρ_0 suppose that $g^{-1} \operatorname{im} \widetilde{\rho_0}\, g = \mathrm{PGL}_2(F)$ and let $H = g\, \mathrm{PSL}_2(F) g^{-1}$. Then $W_\lambda \simeq W_\lambda^*$ over H and

$$(1.18) \qquad H^1(H, W_\lambda) \underset{F}{\otimes} \bar{F} \simeq H^1(g^{-1}Hg, g^{-1}(W_\lambda \underset{F}{\otimes} \bar{F})) = 0.$$

We deduce also that $H^1(\operatorname{im} \rho_0, W_\lambda^*) = 0$.

Finally we consider the case where $F = \mathbf{F}_3$. I am grateful to Taylor for the following argument. First we consider the action of $\mathrm{PSL}_2(\mathbf{F}_3)$ on W_λ explicitly by considering the conjugation action on matrices $\{A \in M_2(\mathbf{F}_3) : \operatorname{trace} A = 0\}$. One sees that no such matrix is fixed by all the elements of order 2, whence

$$H^1(\mathrm{PSL}_2(\mathbf{F}_3), W_\lambda) \simeq H^1(\mathbf{Z}/3, (W_\lambda)^{C_2 \times C_2}) = 0$$

where $C_2 \times C_2$ denotes the normal subgroup of order 4 in $\mathrm{PSL}_2(\mathbf{F}_3) \simeq A_4$. Next we verify that there is a unique copy of A_4 in $\mathrm{PGL}_2(\bar{\mathbf{F}}_3)$ up to conjugation. For suppose that $A, B \in \mathrm{GL}_2(\bar{\mathbf{F}}_3)$ are such that $A^2 = B^2 = I$ with the images of A, B representing distinct nontrivial commuting elements of $\mathrm{PGL}_2(\bar{\mathbf{F}}_3)$. We can choose $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ by a suitable choice of basis, i.e., by a suitable conjugation. Then B is diagonal or antidiagonal as it commutes with A up to a scalar, and as B, A are distinct in $\mathrm{PGL}_2(\bar{\mathbf{F}}_3)$ we have $B = \left(\begin{smallmatrix} 0 & -a^{-1} \\ a & 0 \end{smallmatrix}\right)$ for some a. By conjugating by a diagonal matrix (which does not change A) we can assume that $a = 1$. The group generated by $\{A, B\}$ in $\mathrm{PGL}_2(\mathbf{F}_3)$ is its own centralizer so it has index at most 6 in its normalizer N. Since $N/\langle A, B\rangle \simeq S_3$ there is a unique subgroup of N in which $\langle A, B\rangle$ has index 3 whence the image of the embedding of A_4 in $\mathrm{PGL}_2(\bar{\mathbf{F}}_3)$ is indeed unique (up to conjugation). So arguing as in (1.18) by extending scalars we see that $H^1(\operatorname{im} \rho_0, W_\lambda^*) = 0$ when $F = \mathbf{F}_3$ also. □

The following lemma was pointed out to me by Taylor. It permits most dihedral cases to be covered by the methods of Chapter 3 and [TW].

LEMMA 1.12. *Suppose that ρ_0 is absolutely irreducible and that*

(a) *$\tilde{\rho}_0$ is dihedral (the case where the image is $\mathbf{Z}/2 \times \mathbf{Z}/2$ is allowed),*

(b) $\rho_0|_L$ *is absolutely irreducible where* $L = \mathbf{Q}\left(\sqrt{(-1)^{(p-1)/2}p}\right)$.

Then for any positive integer n *and any irreducible Galois stable subspace* X *of* $W_\lambda \otimes \bar{k}$ *there exists an element* $\sigma \in \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ *such that*

(i) $\tilde{\rho}_0(\sigma) \neq 1$,

(ii) σ *fixes* $\mathbf{Q}(\zeta_{p^n})$,

(iii) σ *has an eigenvalue* 1 *on* X.

Proof. If $\tilde{\rho}_0$ is dihedral then $\rho_0 \otimes \bar{k} = \mathrm{Ind}_H^G \chi$ for some H of index 2 in G, where $G = \mathrm{Gal}(K_1/\mathbf{Q})$. (As before, K_1 is the splitting field of ρ_0.) Here H can be taken as the full inverse image of any of the normal subgroups of index 2 defining the dihedral group. Then $W_\lambda \otimes \bar{k} \simeq \delta \oplus \mathrm{Ind}_H^G(\chi/\chi')$ where δ is the quadratic character $G \to G/H$ and χ' is the conjugate of χ by any element of $G - H$. Note that $\chi \neq \chi'$ since H has nontrivial image in $\mathrm{PGL}_2(\bar{k})$.

To find a σ such that $\delta(\sigma) = 1$ and conditions (i) and (ii) hold, observe that $M(\zeta_{p^n})$ is abelian where M is the quadratic field associated to δ. So conditions (i) and (ii) can be satisfied if $\tilde{\rho}_0$ is non-abelian. If $\tilde{\rho}_0$ is abelian (i.e., the image has the form $\mathbf{Z}/2 \times \mathbf{Z}/2$), then we use hypothesis (b). If $\mathrm{Ind}_H^G(\chi/\chi')$ is reducible over $\bar{k}$ then $W_\lambda \otimes \bar{k}$ is a sum of three distinct quadratic characters, none of which is the quadratic character associated to L, and we can repeat the argument by changing the choice of H for the other two characters. If $X = \mathrm{Ind}_H^G(\chi/\chi') \otimes \bar{k}$ is absolutely irreducible then pick any $\sigma \in G - H$. This satisfies (i) and can be made to satisfy (ii) if (b) holds. Finally, since $\sigma \in G - H$ we see that σ has trace zero and $\sigma^2 = 1$ in its action on X. Thus it has an eigenvalue equal to 1. □

Chapter 2

In this chapter we study the Hecke rings. In the first section we recall some of the well-known properties of these rings and especially the Gorenstein property whose proof is rather technical, depending on a characteristic p version of the q-expansion principle. In the second section we compute the relations between the Hecke rings as the level is augmented. The purpose is to find the change in the η-invariant as the level increases.

In the third section we state the conjecture relating the deformation rings of Chapter 1 and the Hecke rings. Finally we end with the critical step of showing that if the conjecture is true at a minimal level then it is true at all levels. By the results of the appendix the conjecture is equivalent to the

equality of the η-invariant for the Hecke rings and the $\mathfrak{p}/\mathfrak{p}^2$-invariant for the deformation rings. In Chapter 2, Section 2, we compute the change in the η-invariant and in Chapter 1, Section 1, we estimated the change in the $\mathfrak{p}/\mathfrak{p}^2$-invariant.

1. The Gorenstein property

For any positive integer N let $X_1(N) = X_1(N)_{/\mathbf{Q}}$ be the modular curve over $\mathbf{Q}$ corresponding to the group $\Gamma_1(N)$ and let $J_1(N)$ be its Jacobian. Let $\mathbf{T}_1(N)$ be the ring of endomorphisms of $J_1(N)$ which is generated over $\mathbf{Z}$ by the standard Hecke operators $\{T_l = T_{l*}$ for $l \nmid N$, $U_q = U_{q*}$ for $q \mid N$, $\langle a \rangle = \langle a \rangle_*$ for $(a, N) = 1\}$. For precise definitions of these see [MW1, Ch. 2, §5]. In particular if one identifies the cotangent space of $J_1(N)(\mathbf{C})$ with the space of cusp forms of weight 2 on $\Gamma_1(N)$ then the action induced by $\mathbf{T}_1(N)$ is the usual one on cusp forms. We let $\Delta = \{\langle a \rangle : (a, N) = 1\}$.

The group $(\mathbf{Z}/N\mathbf{Z})^*$ acts naturally on $X_1(N)$ via Δ and for any subgroup $H \subseteq (\mathbf{Z}/N\mathbf{Z})^*$ we let $X_H(N) = X_H(N)_{/\mathbf{Q}}$ be the quotient $X_1(N)/H$. Thus for $H = (\mathbf{Z}/N\mathbf{Z})^*$ we have $X_H(N) = X_0(N)$ corresponding to the group $\Gamma_0(N)$. In Section 2 it will sometimes be convenient to assume that H decomposes as a product $H = \prod H_q$ in $(\mathbf{Z}/N\mathbf{Z})^* \simeq \prod (\mathbf{Z}/q^r\mathbf{Z})^*$ where the product is over the distinct prime powers dividing N. We let $J_H(N)$ denote the Jacobian of $X_H(N)$ and note that the above Hecke operators act naturally on $J_H(N)$ also. The ring generated by these Hecke operators is denoted $\mathbf{T}_H(N)$ and sometimes, if H and N are clear from the context, we abbreviate this to $\mathbf{T}$.

Let p be a prime ≥ 3. Let $\mathfrak{m}$ be a maximal ideal of $\mathbf{T} = \mathbf{T}_H(N)$ with $p \in \mathfrak{m}$. Then associated to $\mathfrak{m}$ there is a continuous odd semisimple Galois representation $\rho_\mathfrak{m}$,

$$(2.1) \qquad \rho_\mathfrak{m}\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \to \operatorname{GL}_2(\mathbf{T}/\mathfrak{m})$$

unramified outside Np which satisfies

$$\operatorname{trace} \rho_\mathfrak{m}(\operatorname{Frob} q) = T_q\,, \quad \det \rho_\mathfrak{m}(\operatorname{Frob} q) = \langle q \rangle q$$

for each prime $q \nmid Np$. Here $\operatorname{Frob} q$ denotes a Frobenius at q in $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. The representation $\rho_\mathfrak{m}$ is unique up to isomorphism. If $p \nmid N$ (resp. $p \mid N$) we say that $\mathfrak{m}$ is ordinary if $T_p \notin \mathfrak{m}$ (resp. $U_p \notin \mathfrak{m}$). This implies (cf., for example, theorem 2 of [Wi1]) that for our fixed decomposition group D_p at p,

$$\rho_\mathfrak{m}\Big|_{D_p} \approx \begin{pmatrix} \chi_1 & * \\ 0 & \chi_2 \end{pmatrix}$$

for a suitable choice of basis, with χ_2 unramified and $\chi_2(\operatorname{Frob} p) = T_p \bmod \mathfrak{m}$ (resp. equal to U_p). In particular $\rho_\mathfrak{m}$ is ordinary in the sense of Chapter 1

provided $\chi_1 \neq \chi_2$. We will say that $\mathfrak{m}$ is D_p-distinguished if $\mathfrak{m}$ is ordinary and $\chi_1 \neq \chi_2$. (In practice χ_1 is usually ramified so this imposes no extra condition.) We caution the reader that if $\rho_\mathfrak{m}$ is ordinary in the sense of Chapter 1 then we can only conclude that $\mathfrak{m}$ is D_p-distinguished if $p \nmid N$.

Let $\mathbf{T}_\mathfrak{m}$ denote the completion of $\mathbf{T}$ at $\mathfrak{m}$ so that $\mathbf{T}_\mathfrak{m}$ is a direct factor of the complete semi-local ring $\mathbf{T}_p = \mathbf{T} \otimes \mathbf{Z}_p$. Let $\mathcal{D}$ be the points of the associated $\mathfrak{m}$-divisible group

$$\mathcal{D} = J_H(N)\ (\overline{\mathbf{Q}})_\mathfrak{m} \simeq J_H(N)\ (\overline{\mathbf{Q}})_{p^\infty} \underset{\mathbf{T}_p}{\otimes} \mathbf{T}_\mathfrak{m}\ .$$

It is known that $\hat{\mathcal{D}} = \mathrm{Hom}_{\mathbf{Z}_p}(\mathcal{D}, \mathbf{Q}_p/\mathbf{Z}_p)$ is a rank 2 $\mathbf{T}_\mathfrak{m}$-module, i.e., that $\hat{\mathcal{D}} \underset{\mathbf{Z}_p}{\otimes} \mathbf{Q}_p \simeq (\mathbf{T}_\mathfrak{m} \underset{\mathbf{Z}_p}{\otimes} \mathbf{Q}_p)^2$. Briefly it is enough to show that $H^1(X_H(N), \mathbf{C})$ is free of rank 2 over $\mathbf{T} \otimes \mathbf{C}$ and this reduces to showing that $S_2\ (\Gamma_H(N),\ \mathbf{C})$, the space of cusp forms of weight 2 on $\Gamma_H(N)$, is free of rank 1 over $\mathbf{T} \otimes \mathbf{C}$. One shows then that if $\{f_1, \ldots, f_r\}$ is a complete set of normalized newforms in $S_2(\Gamma_H(N),\ \mathbf{C})$ of levels $m_1, \ldots, m_r$ then if we set $d_i = N/m_i$, the form $f = \Sigma\, f_i(d_i z)$ is a basis vector of $S_2(\Gamma_H(N),\ \mathbf{C})$ as a $\mathbf{T} \otimes \mathbf{C}$-module.

If $\mathfrak{m}$ is ordinary then Theorem 2 of [Wi1], itself a straightforward generalization of Proposition 2 and (11) of [MW2], shows that (for our fixed decomposition group D_p) there is a filtration of $\mathcal{D}$ by Pontrjagin duals of rank 1 $\mathbf{T}_\mathfrak{m}$-modules (in the sense explained above)

(2.2) $$0 \to \mathcal{D}^0 \to \mathcal{D} \to \mathcal{D}^E \to 0$$

where $\mathcal{D}^0$ is stable under D_p and the induced action on $\mathcal{D}^E$ is unramified with $\mathrm{Frob}\, p = U_p$ on it if $p \mid N$ and $\mathrm{Frob}\, p$ equal to the unit root of $x^2 - T_p x + p\langle p\rangle = 0$ in $\mathbf{T}_\mathfrak{m}$ if $p \nmid N$. We can describe $\mathcal{D}^0$ and $\mathcal{D}^E$ as follows. Pick a $\sigma \in I_p$ which induces a generator of $\mathrm{Gal}\,(\mathbf{Q}_p(\zeta_{Np^\infty})/\mathbf{Q}_p(\zeta_{Np}))$. Let $\varepsilon\colon D_p \to \mathbf{Z}_p^\times$ be the cyclotomic character. Then $\mathcal{D}^0 = \ker\,(\sigma - \varepsilon(\sigma))^{\mathrm{div}}$, the kernel being taken inside $\mathcal{D}$ and 'div' meaning the maximal divisible subgroup. Although in [Wi1] this filtration is given only for a factor A_f of $J_1(N)$ it is easy to deduce the result for $J_H(N)$ itself. We note that this filtration is defined without reference to characteristic p and also that if $\mathfrak{m}$ is D_p-distinguished, $\mathcal{D}^0$ (resp. $\mathcal{D}^E$) can be described as the maximal submodule on which $\sigma - \tilde{\chi}_1(\sigma)$ is topologically nilpotent for all $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ (resp. quotient on which $\sigma - \tilde{\chi}_2(\sigma)$ is topologically nilpotent for all $\sigma \in \mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$), where $\tilde{\chi}_i(\sigma)$ is any lifting of $\chi_i(\sigma)$ to $\mathbf{T}_\mathfrak{m}$.

The Weil pairing $\langle\ ,\ \rangle$ on $J_H(N)(\overline{\mathbf{Q}})_{p^M}$ satisfies the relation $\langle t_* x, y\rangle = \langle x, t^* y\rangle$ for any Hecke operator t. It is more convenient to use an adapted pairing defined as follows. Let w_ζ, for ζ a primitive N^{th} root of 1, be the involution of $X_1(N)_{/\mathbf{Q}(\zeta)}$ defined in [MW1, p. 235]. This induces an involution of $X_H(N)_{/\mathbf{Q}(\zeta)}$ also. Then we can define a new pairing $[\ ,\ ]$ by setting (for a

fixed choice of ζ)

$$[x,y]=\langle x,w_\zeta y\rangle. \tag{2.3}$$

Then $[t_*x,y]=[x,t_*y]$ for all Hecke operators t. In particular we obtain an induced pairing on $\mathcal{D}_{pM}$.

The following theorem is the crucial result of this section. It was first proved by Mazur in the case of prime level [Ma2]. It has since been generalized in [Ti1], [Ri1] [M Ri], [Gro] and [E1], but the fundamental argument remains that of [Ma2]. For a summary see [E1, §9]. However some of the cases we need are not covered in these accounts and we will present these here.

THEOREM 2.1. (i) *If* $p\nmid N$ *and* $\rho_{\mathfrak{m}}$ *is irreducible then*

$$J_H(N)(\overline{\mathbf{Q}})[\mathfrak{m}] \simeq (\mathbf{T}/\mathfrak{m})^2.$$

(ii) *If* $p\nmid N$ *and* $\rho_{\mathfrak{m}}$ *is irreducible and* $\mathfrak{m}$ *is* D_p*-distinguished then*

$$J_H(Np)(\overline{\mathbf{Q}})[\mathfrak{m}] \simeq (\mathbf{T}/\mathfrak{m})^2.$$

(*In case* (ii) $\mathfrak{m}$ *is a maximal ideal of* $\mathbf{T}=\mathbf{T}_H(Np)$.)

COROLLARY 1. *In case* (i), $\widehat{J_H(N)(\overline{\mathbf{Q}})}_{\mathfrak{m}}\simeq \mathbf{T}_{\mathfrak{m}}^2$ *and* $\mathrm{Ta}_{\mathfrak{m}}\left(J_H(N)\left(\overline{\mathbf{Q}}\right)\right)\simeq \mathbf{T}_{\mathfrak{m}}^2$.

In case (ii), $\widehat{J_H(Np)(\overline{\mathbf{Q}})}_{\mathfrak{m}}\simeq \mathbf{T}_{\mathfrak{m}}^2$ *and* $\mathrm{Ta}_{\mathfrak{m}}\left(J_H(Np)\left(\overline{\mathbf{Q}}\right)\right)\simeq \mathbf{T}_{\mathfrak{m}}^2$ (*where* $\mathbf{T}_{\mathfrak{m}}=\mathbf{T}_H(Np)_{\mathfrak{m}}$)

COROLLARY 2. *In either of cases* (i) *or* (ii) $\mathbf{T}_{\mathfrak{m}}$ *is a Gorenstein ring.*

In each case the first isomorphisms of Corollary 1 follow from the theorem together with the rank 2 result alluded to previously. Corollary 2 and the second isomorphisms of corollary 1 then follow on applying duality (2.4). (In the proof and in all applications we will only use the notion of a Gorenstein $\mathbf{Z}_p$-algebra as defined in the appendix. For finite flat local $\mathbf{Z}_p$-algebras the notions of Gorenstein ring and Gorenstein $\mathbf{Z}_p$-algebra are the same.) Here $\mathrm{Ta}_{\mathfrak{m}}\left(J_H(N)\left(\overline{\mathbf{Q}}\right)\right)=\mathrm{Ta}_p\left(J_H(N)\left(\overline{\mathbf{Q}}\right)\right)\underset{\mathbf{T}_p}{\otimes}\mathbf{T}_{\mathfrak{m}}$ is the $\mathfrak{m}$-adic Tate module of $J_H(N)$.

We should also point out that although Corollary 1 gives a representation from the $\mathfrak{m}$-adic Tate module

$$\rho=\rho_{\mathbf{T}_{\mathfrak{m}}}:\ \mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})\to \mathrm{GL}_2\left(\mathbf{T}_{\mathfrak{m}}\right)$$

this can be constructed in a much more elementary way. (See [Ca3] for another argument.) For, the representation exists with $\mathbf{T}_{\mathfrak{m}}\otimes\mathbf{Q}$ replacing $\mathbf{T}_{\mathfrak{m}}$ when we use the fact that $\mathrm{Hom}(\mathbf{Q}_p/\mathbf{Z}_p,\mathcal{D})\otimes\mathbf{Q}$ was free of rank 2. A standard argument

using the Eichler-Shimura relations implies that this representation ρ' with values in $\mathrm{GL}_2(\mathbf{T}_\mathfrak{m} \otimes \mathbf{Q})$ has the property that

$$\operatorname{trace} \rho'(\operatorname{Frob} \ell) = T_\ell, \qquad \det \rho'(\operatorname{Frob} \ell) = \ell\langle \ell \rangle$$

for all $\ell \nmid Np$. We can normalize this representation by picking a complex conjugation c and choosing a basis such that $\rho'(c) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, and then by picking a τ for which $\rho'(\tau) = \left(\begin{smallmatrix} a_\tau & b_\tau \\ c_\tau & d_\tau \end{smallmatrix}\right)$ with $b_\tau c_\tau \not\equiv 0(\mathfrak{m})$ and by rescaling the basis so that $b_\tau = 1$. (Note that the explicit description of the traces shows that if $\rho_\mathfrak{m}$ is also normalized so that $\rho_\mathfrak{m}(c) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ then $b_\tau c_\tau \bmod \mathfrak{m} = b_{\tau,\mathfrak{m}} c_{\tau,\mathfrak{m}}$ where $\rho_\mathfrak{m}(\tau) = \left(\begin{smallmatrix} a_{\tau,\mathfrak{m}} & b_{\tau,\mathfrak{m}} \\ c_{\tau,\mathfrak{m}} & d_{\tau,\mathfrak{m}} \end{smallmatrix}\right)$. The existence of a τ such that $b_\tau c_\tau \not\equiv 0(\mathfrak{m})$ comes from the irreducibility of $\rho_\mathfrak{m}$.) With this normalization one checks that ρ' actually takes values in the (closed) subring of $\mathbf{T}_\mathfrak{m}$ generated over $\mathbf{Z}_p$ by the traces. One can even construct the representation directly from the representations in Theorem 0.1 using this ring which is reduced. This is the method of Carayol which requires also the characterization of ρ by the traces and determinants (Theorem 1 of [Ca3]). One can also often interpret the U_q operators in terms of ρ for $q \mid N$ using the $\pi_q \simeq \pi(\sigma_q)$ theorem of Langlands (cf. [Ca1]) and the U_p operator in case (ii) using Theorem 2.1.4 of [Wi1].

Proof (*of theorem*). The important technique for proving such multiplicity-one results is due to Mazur and is based on the q-expansion principle in characteristic p. Since the kernel of $J_H(N)(\overline{\mathbf{Q}}) \to J_1(N)(\overline{\mathbf{Q}})$ is an abelian group on which $\operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ acts through an abelian extension of $\mathbf{Q}$, the intersection with $\ker \mathfrak{m}$ is trivial when $\rho_\mathfrak{m}$ is irreducible. So it is enough to verify the theorem for $J_1(N)$ in part (i) (resp. $J_1(Np)$ in part (ii)). The method for part (i) was developed by Mazur in [Ma2, Ch. II, Prop. 14.2]. It was extended to the case of $\Gamma_0(N)$ in [Ri1, Th. 5.2] which summarizes Mazur's argument. The case of $\Gamma_1(N)$ is similar (cf. [E1, Th. 9.2]).

Now consider case (ii). Let $\Delta_{(p)} = \{\langle a \rangle : a \equiv 1(N)\} \subseteq \Delta$. Let us first assume that $\Delta_{(p)}$ is nontrivial mod $\mathfrak{m}$, i.e., that $\delta - 1 \notin \mathfrak{m}$ for some $\delta \in \Delta_{(p)}$. This case is essentially covered in [Ti1] (and also in [Gro]). We briefly review the argument for use later. Let $K = \mathbf{Q}_p(\zeta_p)$, ζ_p being a primitive p^{th} root of unity, and let $\mathcal{O}$ be the ring of integers of the completion of the maximal unramified extension of K. Using the fact that $\Delta_{(p)}$ is nontrivial mod $\mathfrak{m}$ together with Proposition 4, p. 269 of [MW1] we find that

$$J_1(Np)^{\text{ét}}_{\mathfrak{m}/\mathcal{O}}\,(\overline{\mathbf{F}}_p) \simeq (\operatorname{Pic}^0 \Sigma_1^{\text{ét}} \times \operatorname{Pic}^0 \Sigma_1^{\mu})_\mathfrak{m}\,(\overline{\mathbf{F}}_p)$$

where the notation is taken from [MW1] loc. cit. Here $\Sigma_1^{\text{ét}}$ and Σ_1^{μ} are the two smooth irreducible components of the special fibre of the canonical model of $X_1(Np)_{/\mathcal{O}}$ described in [MW1, Ch. 2]. (The smoothness in this case was proved in [DR].) Also $J_1(Np)^{\text{ét}}_{\mathfrak{m}/\mathcal{O}}$ denotes the canonical étale quotient of the $\mathfrak{m}$-divisible group over $\mathcal{O}$. This makes sense because $J_1(Np)_\mathfrak{m}$ does extend to

a p-divisible group over $\mathcal{O}$ (again by a theorem of Deligne and Rapoport [DR] and because $\Delta_{(p)}$ is nontrivial mod $\mathfrak{m}$). It is ordinary as follows from (2.2) when we use the main theorem of Tate ([Ta]) since $\mathcal{D}^0$ and $\mathcal{D}^E$ clearly correspond to ordinary p-divisible groups.

Now the q-expansion principle implies that $\dim_{\overline{\mathbf{F}}_p} X[\mathfrak{m}'] \leq 1$ where

$$X = \{H^0(\Sigma_1^\mu, \Omega^1) \oplus H^0(\Sigma_1^{\text{ét}}, \Omega^1)\}$$

and $\mathfrak{m}'$ is defined by embedding $\mathbf{T}/\mathfrak{m} \hookrightarrow \overline{\mathbf{F}}_p$ and setting $\mathfrak{m}' = \ker\colon \mathbf{T} \otimes \overline{\mathbf{F}}_p \to \overline{\mathbf{F}}_p$ under the map $t \otimes a \mapsto at \bmod \mathfrak{m}$. Also $\mathbf{T}$ acts on $\operatorname{Pic}^0 \Sigma_1^\mu \times \operatorname{Pic}^0 \Sigma_1^{\text{ét}}$, the abelian variety part of the closed fibre of the Neron model of $J_1(Np)_{/\mathcal{O}}$, and hence also on its cotangent space X. (For a proof that $X[\mathfrak{m}']$ is at most one-dimensional, which is readily adapted to this case, see Lemma 2.2 below. For similar versions in slightly simpler contexts see [Wi3, §6] or [Gro, §12].) Then the Cartier map induces an injection (cf. Prop. 6.5 of [Wi3])

$$\delta\colon \{\operatorname{Pic}^0 \Sigma_1^\mu \times \operatorname{Pic}^0 \Sigma_1^{\text{ét}}\}[p](\overline{\mathbf{F}}_p) \underset{\mathbf{F}_p}{\otimes} \overline{\mathbf{F}}_p \hookrightarrow X.$$

The composite $\delta \circ w_\zeta$ can be checked to be Hecke invariant (cf. Prop. 6.5 of [Wi3]. In checking the compatibility for U_p use the formulas of Theorem 5.3 of [Wi3] but note the correction in [MW1, p. 188].) It follows that

$$J_1(Np)_{\mathfrak{m}/\mathcal{O}}\,(\overline{\mathbf{F}}_p)\,[\mathfrak{m}] \;\simeq\; \mathbf{T}/\mathfrak{m}$$

as a $\mathbf{T}$-module. This shows that if $\hat{H}$ is the Pontrjagin dual of $H = J_1(Np)_{\mathfrak{m}/\mathcal{O}}(\overline{\mathbf{F}}_p)$ then $\hat{H} \simeq \mathbf{T}_\mathfrak{m}$ since $\hat{H}/\mathfrak{m} \simeq \mathbf{T}/\mathfrak{m}$. Thus

$$J_1(Np)_{\mathfrak{m}/\mathcal{O}}(\overline{\mathbf{F}}_p)\,[p] \xrightarrow{\sim} \operatorname{Hom}(\mathbf{T}_\mathfrak{m}/p,\ \mathbf{Z}/p\mathbf{Z}).$$

Now our assumption that $\mathfrak{m}$ is D_p-distinguished enables us to identify

$$\mathcal{D}^0 = J_1(Np)^0_{\mathfrak{m}/\mathcal{O}}(\overline{\mathbf{Q}}_p) \quad , \quad \mathcal{D}^E = J_1(Np)^{\text{ét}}_{\mathfrak{m}/\mathcal{O}}(\overline{\mathbf{Q}}_p).$$

For the groups on the right are unramified and those on the left are dual to groups where inertia acts via a character of finite order (duality with respect to $\operatorname{Hom}(\ ,\mathbf{Q}_p/\mathbf{Z}_p(1))$). So

$$\mathcal{D}^0[p] \xrightarrow{\sim} \mathbf{T}_\mathfrak{m}/p, \quad \mathcal{D}^E[p] \xrightarrow{\sim} \operatorname{Hom}(\mathbf{T}_\mathfrak{m}/p,\ \mathbf{Z}/p\mathbf{Z})$$

as $\mathbf{T}_\mathfrak{m}$-modules, the former following from the latter when we use duality under the pairing $[\ ,]$. In particular as $\mathfrak{m}$ is D_p-distinguished,

$$\mathcal{D}[p] \simeq \mathbf{T}_\mathfrak{m}/p \,\oplus\, \operatorname{Hom}(\mathbf{T}_\mathfrak{m}/p,\ \mathbf{Z}/p\mathbf{Z}). \tag{2.4}$$

We now use an argument of Tilouine [Til]. We pick a complex conjugation τ. This has distinct eigenvalues ± 1 on $\rho_\mathfrak{m}$ so we may decompose $\mathcal{D}[p]$ into eigenspaces for τ:

$$\mathcal{D}[p] = \mathcal{D}[p]^+ \oplus \mathcal{D}[p]^-.$$

Since $\mathbf{T}_\mathfrak{m}/p$ and Hom $(\mathbf{T}_\mathfrak{m}/p,\ \mathbf{Z}/p\mathbf{Z})$ are both indecomposable Hecke-modules, by the Krull-Schmidt theorem this decomposition has factors which are isomorphic to those in (2.4) up to order. So in the decomposition

$$\mathcal{D}[\mathfrak{m}] = \mathcal{D}[\mathfrak{m}]^+ \oplus \mathcal{D}[\mathfrak{m}]^-$$

one of the eigenspaces is isomorphic to $\mathbf{T}/\mathfrak{m}$ and the other to $(\mathbf{T}_\mathfrak{m}/p)[\mathfrak{m}]$. But since $\rho_\mathfrak{m}$ is irreducible it is easy to see by considering $\mathcal{D}[\mathfrak{m}] \oplus \mathrm{Hom}(\mathcal{D}[\mathfrak{m}], \det \rho_\mathfrak{m})$ that τ has the same number of eigenvalues equal to $+1$ as equal to -1 in $\mathcal{D}[\mathfrak{m}]$, whence $\#(\mathbf{T}_\mathfrak{m}/p)\,[\mathfrak{m}] = \#(\mathbf{T}/\mathfrak{m})$. This shows that $\mathcal{D}[\mathfrak{m}]^+ \xrightarrow{\sim} \mathcal{D}[\mathfrak{m}]^- \simeq \mathbf{T}/\mathfrak{m}$ as required.

Now we consider the case where $\Delta_{(p)}$ is trivial mod $\mathfrak{m}$. This case was treated (but only for the group $\Gamma_0(Np)$ and $\rho_\mathfrak{m}$ 'new' at p—the crucial restriction being the last one) in [M Ri]. Let $X_1(N,p)_{/\mathbf{Q}}$ be the modular curve corresponding to $\Gamma_1(N) \cap \Gamma_0(p)$ and let $J_1(N,p)$ be its Jacobian. Then since the composite of natural maps $J_1(N,p) \to J_1(Np) \to J_1(N,p)$ is multiplication by an integer prime to p and since $\Delta_{(p)}$ is trivial mod $\mathfrak{m}$ we see that

$$J_1(N,p)_\mathfrak{m}(\overline{\mathbf{Q}}) \simeq J_1(Np)_\mathfrak{m}(\overline{\mathbf{Q}}).$$

It will be enough then to use $J_1(N,p)$, and the corresponding ring $\mathbf{T}$ and ideal $\mathfrak{m}$.

The curve $X_1(N,p)$ has a canonical model $X_1(N,p)_{/\mathbf{Z}_p}$ which over $\overline{\mathbf{F}}_p$ consists of two smooth curves $\Sigma^{\text{ét}}$ and Σ^μ intersecting transversally at the supersingular points (again this is a theorem of Deligne and Rapoport; cf. [DR, Ch. 6, Th. 6.9], [KM] or [MW1] for more details). We will use the models described in [MW1, Ch. II] and in particular the cusp ∞ will lie on Σ^μ. Let Ω denote the sheaf of regular differentials on $X_1(N,p)_{/\mathbf{F}_p}$ (cf. [DR, Ch. 1 §2], [M Ri, §7]). Over $\overline{\mathbf{F}}_p$, since $X_1(N,p)_{/\overline{\mathbf{F}}_p}$ has ordinary double point singularities, the differentials may be identified with the meromorphic differentials on the normalization $\widetilde{X_1(N,p)}_{/\overline{\mathbf{F}}_p} = \Sigma^{\text{ét}} \cup \Sigma^\mu$ which have at most simple poles at the supersingular points (the intersection points of the two components) and satisfy $\mathrm{res}_{x_1} + \mathrm{res}_{x_2} = 0$ if x_1 and x_2 are the two points above such a supersingular point. We need the following lemma:

LEMMA 2.2. $\dim_{\mathbf{T}/\mathfrak{m}} H^0(X_1(N,p)_{/\mathbf{F}_p}, \Omega)\,[\mathfrak{m}] = 1.$

Proof. First we remark that the action of the Hecke operator U_p here is most conveniently defined using an extension from characteristic zero. This is explained below. We will first show that $\dim_{\mathbf{T}/\mathfrak{m}} H^0(X_1(N,p)_{/\mathbf{F}_p}, \Omega)\,[\mathfrak{m}] \leq 1$, this being the essential step. If we embed $\mathbf{T}/\mathfrak{m} \hookrightarrow \overline{\mathbf{F}}_p$ and then set $\mathfrak{m}' = \ker\colon \mathbf{T} \otimes \overline{\mathbf{F}}_p \longrightarrow \overline{\mathbf{F}}_p$ (the map given by $t \otimes a \mapsto at \bmod \mathfrak{m}$) then it is enough to show that $\dim_{\overline{\mathbf{F}}_p} H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p}, \Omega)[\mathfrak{m}'] \leq 1$. First we will suppose

that there is no nonzero holomorphic differential in $H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p},\Omega)\,[\mathfrak{m}']$, i.e., no differential form which pulls back to *holomorphic* differentials on $\Sigma^{\text{ét}}$ and Σ^μ. Then if ω_1 and ω_2 are two differentials in $H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p},\Omega)\,[\mathfrak{m}']$, the q-expansion principle shows that $\mu\omega_1-\lambda\omega_2$ has zero q-expansion at ∞ for some pair $(\mu,\lambda)\neq(0,0)$ in $\overline{\mathbf{F}}_p^2$ and thus is zero on Σ^μ. As $\mu\omega_1-\lambda\omega_2=0$ on Σ^μ it is holomorphic on $\Sigma^{\text{ét}}$. By our hypothesis it would then be zero which shows that ω_1 and ω_2 are linearly dependent.

This use of the q-expansion principle in characteristic p is crucial and due to Mazur [Ma2]. The point is simply that all the coefficients in the q-expansion are determined by elementary formulae from the coefficient of q provided that ω is an eigenform for all the Hecke operators. The formulae for the action of these operators in characteristic p follow from the formulae in characteristic zero. To see this formally (especially for the U_p operator) one checks first that $H^0(X_1(N,p)_{/\mathbf{Z}_p},\Omega)$, where Ω denotes the sheaf of regular differentials on $X_1(N,p)_{/\mathbf{Z}_p}$, behaves well under the base changes $\mathbf{Z}_p\to\overline{\mathbf{F}}_p$ and $\mathbf{Z}_p\to\overline{\mathbf{Q}}_p$; cf. [Ma2, §II.3] or [Wi3, Prop. 6.1]. The action of the Hecke operators on $J_1(N,p)$ induces an action on the connected component of the Neron model of $J_1(N,p)_{/\mathbf{Q}_p}$, so also on its tangent space and cotangent space. By Grothendieck duality the cotangent space is isomorphic to $H^0(X_1(N,p)_{/\mathbf{Z}_p},\Omega)$; see (2.5) below. (For a summary of the duality statements used in this context, see [Ma2, §II.3]. For explicit duality over fields see [AK, Ch. VIII].) This then defines an action of the Hecke operators on this group. To check that over $\overline{\mathbf{Q}}_p$ this gives the standard action one uses the commutativity of the diagram after Proposition 2.2 in [Mi1].

Now assume that there is a nonzero holomorphic differential in

$$H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p},\,\Omega)\,[\mathfrak{m}'].$$

We claim that the space of holomorphic differentials then has dimension 1 and that any such differential $\omega\neq0$ is actually nonzero on Σ^μ. The dimension claim follows from the second assertion by using the q-expansion principle. To prove that $\omega\neq0$ on Σ^μ we use the formula

$$U_{p*}(x,y)=(Fx,\,y')$$

for $(x,y)\in(\operatorname{Pic}^0\Sigma^{\text{ét}}\times\operatorname{Pic}^0\Sigma^\mu)(\overline{\mathbf{F}}_p)$, where F denotes the Frobenius endomorphism. The value of y' will not be needed. This formula is a variant on the second part of Theorem 5.3 of [Wi3] where the corresponding result is proved for $X_1(Np)$. (A correction to the first part of Theorem 5.3 was noted in [MW1, p. 188].) One checks then that the action of U_p on $X_0=H^0(\Sigma^\mu,\Omega^1)\oplus H^0(\Sigma^{\text{ét}},\Omega^1)$ viewed as a subspace of $H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p},\Omega)$ is the same as the action on X_0 viewed as the cotangent space of $\operatorname{Pic}^0\Sigma^\mu\times\operatorname{Pic}^0\Sigma^{\text{ét}}$. From this we see that if $\omega=0$ on Σ^μ then $U_p\omega=0$ on $\Sigma^{\text{ét}}$. But U_p

acts as a nonzero scalar which gives a contradiction if $\omega \neq 0$. We can thus assume that the space of $\mathfrak{m}'$-torsion holomorphic differentials has dimension 1 and is generated by ω. So if ω_2 is now any differential in $H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p},\Omega)[\mathfrak{m}']$ then $\omega_2-\lambda\omega$ has zero q-expansion at ∞ for some choice of λ. Then $\omega_2-\lambda\omega=0$ on Σ^μ whence $\omega_2-\lambda\omega$ is holomorphic and so $\omega_2=\lambda\omega$. We have now shown in general that $\dim(H^0(X_1(N,p)_{/\overline{\mathbf{F}}_p},\Omega)[\mathfrak{m}'])\leq 1$.

The singularities of $X_1(N,p)_{/\mathbf{Z}_p}$ at the supersingular points are formally isomorphic over $\widehat{\mathbf{Z}_p^{\mathrm{unr}}}$ to $\widehat{\mathbf{Z}_p^{\mathrm{unr}}}[[X,Y]]/(XY-p^k)$ with $k=1$, 2 or 3 (cf. [DR, Ch. 6, Th. 6.9]). If we consider a minimal regular resolution $M_1(N,p)_{/\mathbf{Z}_p}$ then $H^0(M_1(N,p)_{/\mathbf{F}_p},\Omega)\simeq H^0(X_1(N,p)_{/\mathbf{F}_p},\Omega)$ (see the argument in [Ma2, Prop. 3.4]), and a similar isomorphism holds for $H^0(M_1(N,p)_{/\mathbf{Z}_p},\Omega)$.

As $M_1(N,p)_{/\mathbf{Z}_p}$ is regular, a theorem of Raynaud [Ray2] says that the connected component of the Neron model of $J_1(N,p)_{/\mathbf{Q}_p}$ is $J_1(N,p)^0_{/\mathbf{Z}_p}\simeq \mathrm{Pic}^0(M_1(N,p)_{/\mathbf{Z}_p})$. Taking tangent spaces at the origin, we obtain

$$(2.5)\qquad \mathrm{Tan}(J_1(N,p)^0_{/\mathbf{Z}_p})\simeq H^1(M_1(N,p)_{/\mathbf{Z}_p},\mathcal{O}_{M_1(N,p)}).$$

Reducing both sides mod p and applying Grothendieck duality we get an isomorphism

$$(2.6)\qquad \mathrm{Tan}(J_1(N,p)^0_{/\mathbf{F}_p})\xrightarrow{\sim}\mathrm{Hom}(H^0(X_1(N,p)_{/\mathbf{F}_p},\Omega),\mathbf{F}_p).$$

(To justify the reduction in detail see the arguments in [Ma2, §II. 3]). Since $\mathrm{Tan}(J_1(N,p)^0_{/\mathbf{Z}_p})$ is a faithful $\mathbf{T}\otimes\mathbf{Z}_p$-module it follows that

$$H^0(X_1(N,p)_{/\mathbf{F}_p},\Omega)[\mathfrak{m}]$$

is nonzero. This completes the proof of the lemma. □

To complete the proof of the theorem we choose an abelian subvariety A of $J_1(N,p)$ with multiplicative reduction at p. Specifically let A be the connected part of the kernel of $J_1(N,p)\to J_1(N)\times J_1(N)$ under the natural map $\hat{\varphi}$ described in Section 2 (see (2.10)). Then we have an exact sequence

$$0\to A\to J_1(N,p)\to B\to 0$$

and $J_1(N,p)$ has semistable reduction over $\mathbf{Q}_p$ and B has good reduction. By Proposition 1.3 of [Ma3] the corresponding sequence of connected group schemes

$$0\to A[p]^0_{/\mathbf{Z}_p}\to J_1(N,p)[p]^0_{/\mathbf{Z}_p}\to B[p]^0_{/\mathbf{Z}_p}\to 0$$

is also exact, and by Corollary 1.1 of the same proposition the corresponding sequence of tangent spaces of Neron models is exact. Using this we may check that the natural map

$$(2.7)\qquad \mathrm{Tan}(J_1(N,p)[p]^t_{/\overline{\mathbf{F}}_p})\underset{\mathbf{T}_p}{\otimes}\mathbf{T}_\mathfrak{m}\to \mathrm{Tan}(J_1(N,p)_{/\overline{\mathbf{F}}_p})\underset{\mathbf{T}_p}{\otimes}\mathbf{T}_\mathfrak{m}$$

is an isomorphism, where t denotes the maximal multiplicative-type subgroup scheme (cf. [Ma3, §1]). For it is enough to check such a relation on A and B separately and on B it is true because the $\mathfrak{m}$-divisible group is ordinary. This follows from (2.2) by the theorem of Tate [Ta] as before.

Now (2.6) together with the lemma shows that

$$\mathrm{Tan}(J_1(N,p))_{/\mathbf{Z}_p} \underset{\mathbf{T}_p}{\otimes} \mathbf{T}_\mathfrak{m} \simeq \mathbf{T}_\mathfrak{m}.$$

We claim that (2.7) together with this implies that as $\mathbf{T}_\mathfrak{m}$-modules

$$V := J_1(N,p)\,[p]^t(\overline{\mathbf{Q}}_p)_\mathfrak{m} \simeq (\mathbf{T}_\mathfrak{m}/p).$$

To see this it is sufficient to exhibit an isomorphism of $\overline{\mathbf{F}}_p$-vector spaces

$$\mathrm{Tan}(G_{/\overline{\mathbf{F}}_p}) \simeq G(\overline{\mathbf{Q}}_p) \underset{\mathbf{F}_p}{\otimes} \overline{\mathbf{F}}_p \tag{2.8}$$

for any multiplicative-type group scheme (finite and flat) $G_{/\mathbf{Z}_p}$ which is killed by p and moreover to give such an isomorphism that respects the action of endomorphisms of $G_{/\mathbf{Z}_p}$. To obtain such an isomorphism observe that we have isomorphisms

$$\begin{aligned} \mathrm{Hom}_{\overline{\mathbf{Q}}_p}(\boldsymbol{\mu}_p, G) \underset{\mathbf{F}_p}{\otimes} \overline{\mathbf{F}}_p &\simeq \mathrm{Hom}_{\overline{\mathbf{F}}_p}(\boldsymbol{\mu}_p, G) \underset{\mathbf{F}_p}{\otimes} \overline{\mathbf{F}}_p \\ &\simeq \mathrm{Hom}\left(\mathrm{Tan}(\boldsymbol{\mu}_{p/\overline{\mathbf{F}}_p}), \mathrm{Tan}(G_{/\overline{\mathbf{F}}_p})\right) \end{aligned} \tag{2.9}$$

where $\mathrm{Hom}_{\overline{\mathbf{Q}}_p}$ denotes homomorphisms of the group schemes viewed over $\overline{\mathbf{Q}}_p$ and similarly for $\mathrm{Hom}_{\overline{\mathbf{F}}_p}$. The second isomorphism can be checked by reducing to the case $G = \boldsymbol{\mu}_p$. Now picking a primitive p^{th} root of unity we can identify the left-hand term in (2.9) with $G(\overline{\mathbf{Q}}_p) \underset{\mathbf{F}_p}{\otimes} \overline{\mathbf{F}}_p$. Picking an isomorphism of $\mathrm{Tan}(\boldsymbol{\mu}_{p/\overline{\mathbf{F}}_p})$ with $\overline{\mathbf{F}}_p$ we can identify the last term in (2.9) with $\mathrm{Tan}(G_{/\overline{\mathbf{F}}_p})$. Thus after these choices are made we have an isomorphism in (2.8) which respects the action of endomorphisms of G.

On the other hand the action of $\mathrm{Gal}(\overline{\mathbf{Q}}_p/\mathbf{Q}_p)$ on V is ramified on every subquotient, so $V \subseteq \mathcal{D}^0[p]$. (Note that our assumption that $\Delta_{(p)}$ is trivial $\bmod\, \mathfrak{m}$ implies that the action on $\mathcal{D}^0[p]$ is ramified on every subquotient and on $\mathcal{D}^E[p]$ is unramified on every subquotient.) By again examining A and B separately we see that in fact $V = \mathcal{D}^0[p]$. For A we note that $A[p]/A[p]^t$ is unramified because it is dual to $\widehat{A}[p]^t$ where $\widehat{A}$ is the dual abelian variety. We can now proceed as we did in the case where $\Delta_{(p)}$ was nontrivial $\bmod\, \mathfrak{m}$. $\square$

2. Congruences between Hecke rings

Suppose that q is a prime not dividing N. Let $\Gamma_1(N,q) = \Gamma_1(N) \cap \Gamma_0(q)$ and let $X_1(N,q) = X_1(N,q)_{/\mathbf{Q}}$ be the corresponding curve. The two natural maps $X_1(N,q) \to X_1(N)$ induced by the maps $z \to z$ and $z \to qz$ on the upper half plane permit us to define a map $J_1(N) \times J_1(N) \to J_1(N,q)$. Using a theorem of Ihara, Ribet shows that this map is injective (cf. [Ri2, Cor. 4.2]). Thus we can define φ by

$$(2.10) \qquad 0 \to J_1(N) \times J_1(N) \xrightarrow{\varphi} J_1(N,q).$$

Dualizing, we define B by

$$0 \to B \xrightarrow{\psi} J_1(N,q) \xrightarrow{\hat{\varphi}} J_1(N) \times J_1(N) \to 0.$$

Let $\mathbf{T}_1(N,q)$ be the ring of endomorphisms of $J_1(N,q)$ generated by the standard Hecke operators $\{T_{l*}$ for $l \nmid Nq$, U_{l*} for $l \mid Nq$, $\langle a \rangle = \langle a \rangle_*$ for $(a, Nq) = 1\}$. One can check that U_q preserves B either by an explicit calculation or by noting that B is the maximal abelian subvariety of $J_1(N,q)$ with multiplicative reduction at q. We set $J_2 = J_1(N) \times J_1(N)$.

More generally, one can consider $J_H(N)$ and $J_H(N,q)$ in place of $J_1(N)$ and $J_1(N,q)$ (where $J_H(N,q)$ corresponds to $X_1(N,q)/H$) and we write $\mathbf{T}_H(N)$ and $\mathbf{T}_H(N,q)$ for the associated Hecke rings. In this case the corresponding map φ may have a kernel. However since the kernel of $J_H(N) \to J_1(N)$ does not meet $\ker \mathfrak{m}$ for any maximal ideal $\mathfrak{m}$ whose associated $\rho_\mathfrak{m}$ is irreducible, the above sequences remain exact if we restrict to $\mathfrak{m}^{(q)}$-divisible groups, $\mathfrak{m}^{(q)}$ being the maximal ideal associated to $\mathfrak{m}$ of the ring $\mathbf{T}_H^{(q)}(N,q)$ generated by the standard Hecke operators but omitting U_q. With this minor modification the proofs of the results below for $H \neq 1$ follow from the cases of full level. We will use the same notation in the general case. Thus φ is the map $J_2 = J_H(N)^2 \to J_H(N,q)$ induced by $z \to z$ and $z \to qz$ on the two factors, and $B = \ker \hat{\varphi}$. (B will not be an abelian variety in general.)

The following lemma is a straightforward generalization of a lemma of Ribet ([Ri2]). Let n_q be an integer satisfying $n_q \equiv q(N)$ and $n_q \equiv 1(q)$, and write $\langle q \rangle = \langle n_q \rangle \in \mathbf{T}_H(Nq)$.

LEMMA 2.3 (Ribet). *$\psi(B) \cap \varphi(J_2)_{\mathfrak{m}^{(q)}} = \varphi(J_2)\ [U_q^2 - \langle q \rangle]_{\mathfrak{m}^{(q)}}$ for irreducible $\rho_\mathfrak{m}$.*

Proof. The left-hand side is $(\operatorname{im}\varphi \cap \ker\hat{\varphi})$, so we compute $\varphi^{-1}(\operatorname{im}\varphi \cap \ker\hat{\varphi}) = \ker(\hat{\varphi}\circ\varphi)$.

An explicit calculation shows that

$$\hat{\varphi}\circ\varphi = \begin{bmatrix} q+1 & T_q \\ T_q^* & q+1 \end{bmatrix} \text{ on } J_2$$

where $T_q^* = T_q \cdot \langle q \rangle^{-1}$. The matrix action here is on the left. We also find that on J_2

$$U_q \circ \varphi = \varphi \circ \begin{bmatrix} 0 & -\langle q \rangle \\ q & T_q \end{bmatrix}, \tag{2.11}$$

whence

$$(U_q^2 - \langle q \rangle) \circ \varphi \quad = \quad \varphi \circ \begin{bmatrix} -\langle q \rangle & 0 \\ T_q & -\langle q \rangle \end{bmatrix} \circ (\hat{\varphi} \circ \varphi). \qquad \square$$

Now suppose that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(N)$, $p \in \mathfrak{m}$ and $\rho_\mathfrak{m}$ is irreducible. We will now give a slightly stronger result than that given in the lemma in the special case $q = p$. (The case $q \neq p$ we will also strengthen but we will do this separately.) Assume then that $p \nmid N$ and $T_p \notin \mathfrak{m}$. Let a_p be the unit root of $x^2 - T_p x + p\langle p \rangle = 0$ in $\mathbf{T}_H(N)_\mathfrak{m}$. We first define a maximal ideal $\mathfrak{m}_p$ of $\mathbf{T}_H(N,p)$ with the same associated representation as $\mathfrak{m}$. To do this consider the ring

$$S_1 = \mathbf{T}_H(N)[U_1]/(U_1^2 - T_p U_1 + p\langle p \rangle) \subseteq \operatorname{End}(J_H(N)^2)$$

where U_1 is the endomorphism of $J_H(N)^2$ given by the matrix

$$\begin{bmatrix} T_p & -\langle p \rangle \\ p & 0 \end{bmatrix}.$$

It is thus compatible with the action of U_p on $J_H(N,p)$ when compared using $\hat{\varphi}$. Now $\mathfrak{m}_1 = (\mathfrak{m}, U_1 - \widetilde{a_p})$ is a maximal ideal of S_1 where $\widetilde{a_p}$ is any element of $\mathbf{T}_H(N)$ representing the class $\bar{a}_p \in \mathbf{T}_H(N)_\mathfrak{m}/\mathfrak{m} \simeq \mathbf{T}_H(N)/\mathfrak{m}$. Moreover $S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_\mathfrak{m}$ and we let $\mathfrak{m}_p$ be the inverse image of $\mathfrak{m}_1$ in $\mathbf{T}_H(N,p)$ under the natural map $\mathbf{T}_H(N,p) \to S_1$. One checks that $\mathfrak{m}_p$ is D_p-distinguished. For any standard Hecke operator t except U_p (i.e., $t = T_l$, $U_{q'}$ for $q' \neq p$ or $\langle a \rangle$) the image of t is t. The image of U_p is U_1.

We need to check that the induced map

$$\alpha : \mathbf{T}_H(N,p)_{\mathfrak{m}_p} \longrightarrow S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_\mathfrak{m}$$

is surjective. The only problem is to show that T_p is in the image. In the present context one can prove this using the surjectivity of $\hat{\varphi}$ in (2.12) and using the fact that the Tate-modules in the range and domain of $\hat{\varphi}$ are free of rank 2 by Corollary 1 to Theorem 2.1. The result then follows from Nakayama's lemma as one deduces easily that $\mathbf{T}_H(N)_\mathfrak{m}$ is a cyclic $\mathbf{T}_H(N,p)_{\mathfrak{m}_p}$-module. This argument was suggested by Diamond. A second argument using representations can be found at the end of Proposition 2.15. We will now give a third and more direct proof due to Ribet (cf. [Ri4, Prop. 2]) but found independently and shown to us by Diamond.

For the following lemma we let $\mathbf{T}^M$, for an integer M, denote the subring of $\operatorname{End}\left(S_2(\Gamma_1(N))\right)$ generated by the Hecke operators T_n for positive integers n relatively prime to M. Here $S_2\left(\Gamma_1(N)\right)$ denotes the vector space of weight 2 cusp forms on $\Gamma_1(N)$. Write $\mathbf{T}$ for $\mathbf{T}^1$. It will be enough to show that T_p is a redundant operator in $\mathbf{T}^1$, i.e., that $\mathbf{T}^p = \mathbf{T}$. The result for $\mathbf{T}_H(N)_\mathfrak{m}$ then follows.

LEMMA (Ribet). *Suppose that $(M,N) = 1$. If M is odd then $\mathbf{T}^M = \mathbf{T}$. If M is even then $\mathbf{T}^M$ has finite index in $\mathbf{T}$ equal to a power of* 2.

As the rings are finitely generated free $\mathbf{Z}$-modules, it suffices to prove that $\mathbf{T}^M \otimes \mathbf{F}_l \to \mathbf{T} \otimes \mathbf{F}_l$ is surjective unless l and M are both even. The claim follows from

1. $\mathbf{T}^M \otimes \mathbf{F}_l \to \mathbf{T}^{M/p} \otimes \mathbf{F}_l$ is surjective if $p \mid M$ and $p \nmid lN$.

2. $\mathbf{T}^l \otimes \mathbf{F}_l \to \mathbf{T} \otimes \mathbf{F}_l$ is surjective if $l \nmid 2N$.

Proof of 1. Let A denote the Tate module $\mathrm{Ta}_l(J_1(N))$. Then $R = \mathbf{T}^{M/p} \otimes \mathbf{Z}_l$ acts faithfully on A. Let $R' = (R \underset{\mathbf{Z}_l}{\otimes} \mathbf{Q}_l) \cap \operatorname{End}_{\mathbf{Z}_l} A$ and choose d so that $l^d R' \subset lR$. Consider the $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$-module $B = J_1(N)[l^d] \times \mu_{Nl^d}$. By Čebotarev density, there is a prime q not dividing MNl so that $\operatorname{Frob} p = \operatorname{Frob} q$ on B. Using the fact that $T_r = \operatorname{Frob} r + \langle r \rangle r (\operatorname{Frob} r)^{-1}$ on A for $r = p$ and $r = q$, we see that $T_p = T_q$ on $J_1(N)[l^d]$. It follows that $T_p - T_q$ is in $l^d \operatorname{End}_{\mathbf{Z}_l} A$ and therefore in $l^d R' \subset lR$. □

Proof of 2. Let S be the set of cusp forms in $S_2(\Gamma_1(N))$ whose q-expansions at ∞ have coefficients in $\mathbf{Z}$. Recall that $S_2(\Gamma_1(N)) = S \otimes \mathbf{C}$ and that S is stable under the action of $\mathbf{T}$ (cf. [Sh1, Ch. 3] and [Hi4, §4]). The pairing $\mathbf{T} \otimes S \to \mathbf{Z}$ defined by $T \otimes f \mapsto a_1(Tf)$ is easily checked to induce an isomorphism of $\mathbf{T}$-modules

$$S \cong \operatorname{Hom}_{\mathbf{Z}}(\mathbf{T}, \mathbf{Z}).$$

The surjectivity of $\mathbf{T}^l / l\mathbf{T}^l \to \mathbf{T}/l\mathbf{T}$ is equivalent to the injectivity of the dual map

$$\operatorname{Hom}(\mathbf{T}, \mathbf{F}_l) \to \operatorname{Hom}(\mathbf{T}^l, \mathbf{F}_l).$$

Now use the isomorphism $S/lS \cong \operatorname{Hom}(\mathbf{T}, \mathbf{F}_l)$ and note that if f is in the kernel of $S \to \operatorname{Hom}(\mathbf{T}^l, \mathbf{F}_l)$, then $a_n(f) = a_1(T_n f)$ is divisible by l for all n prime to l. But then the mod l form defined by f is in the kernel of the operator $q\frac{d}{dq}$, and is therefore trivial if l is odd. (See Corollary 5 of the main theorem of [Ka].) Therefore f is in lS.

Remark. The argument does not prove that $\mathbf{T}^{Md} = \mathbf{T}^d$ if $(d, N) \neq 1$.

We now return to the assumptions that $\rho_\mathfrak{m}$ is irreducible, $p \nmid N$ and $T_p \notin \mathfrak{m}$. Next we define a principal ideal (Δ_p) of $\mathbf{T}_H(N)_\mathfrak{m}$ as follows. Since $\mathbf{T}_H(N,p)_{\mathfrak{m}_p}$ and $\mathbf{T}_H(N)_\mathfrak{m}$ are both Gorenstein rings (by Corollary 2 of Theorem 2.1) we can define an adjoint $\hat{\alpha}$ to

$$\alpha\colon \mathbf{T}_H(N,p)_{\mathfrak{m}_p} \longrightarrow S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_\mathfrak{m}$$

in the manner described in the appendix and we set $\Delta_p = (\alpha \circ \hat{\alpha})(1)$. Then (Δ_p) is independent of the choice of (Hecke-module) pairings on $\mathbf{T}_H(N,p)_{\mathfrak{m}_p}$ and $\mathbf{T}_H(N)_\mathfrak{m}$. It is equal to the ideal generated by any composite map

$$\mathbf{T}_H(N)_\mathfrak{m} \xrightarrow{\beta} \mathbf{T}_H(N,p)_{\mathfrak{m}_p} \xrightarrow{\alpha} \mathbf{T}_H(N)_\mathfrak{m}$$

provided that β is an injective map of $\mathbf{T}_H(N,p)_{\mathfrak{m}_p}$-modules with $\mathbf{Z}_p$ torsion-free cokernel. (The module structure on $\mathbf{T}_H(N)_\mathfrak{m}$ is defined via α.)

PROPOSITION 2.4. *Assume that $\mathfrak{m}$ is D_p-distinguished and that $\rho_\mathfrak{m}$ is irreducible of level N with $p \nmid N$. Then*

$$(\Delta_p) = \left(T_p^2 - \langle p\rangle(1+p)^2\right) = (a_p^2 - \langle p\rangle).$$

Proof. Consider the maps on p-adic Tate-modules induced by φ and $\hat{\varphi}$:

$$\mathrm{Ta}_p\left(J_H(N)^2\right) \xrightarrow{\varphi} \mathrm{Ta}_p\left(J_H(N,p)\right) \xrightarrow{\widehat{\varphi}} \mathrm{Ta}_p\left(J_H(N)^2\right).$$

These maps commute with the standard Hecke operators with the exception of T_p or U_p (which are not even defined on all the terms). We define

$$S_2 = \mathbf{T}_H(N)[U_2]\big/(U_2^2 - T_p\,U_2 + p\langle p\rangle) \subseteq \mathrm{End}\left(J_H(N)^2\right)$$

where U_2 is the endomorphism of $J_H(N)^2$ defined by $\left(\begin{smallmatrix}0 & -\langle p\rangle\\ p & T_p\end{smallmatrix}\right)$. It satisfies $\varphi U_2 = U_p\varphi$. Again $\mathfrak{m}_2 = (\mathfrak{m},\ U_2 - \widetilde{a_p})$ is a maximal ideal of S_2 and we have, on restricting to the $\mathfrak{m}_1$, $\mathfrak{m}_p$ and $\mathfrak{m}_2$-adic Tate-modules:

$$\begin{array}{ccccc}
\mathrm{Ta}_{\mathfrak{m}_2}\left(J_H(N)^2\right) & \xrightarrow{\varphi} & \mathrm{Ta}_{\mathfrak{m}_p}\left(J_H(N,p)\right) & \xrightarrow{\widehat{\varphi}} & \mathrm{Ta}_{\mathfrak{m}_1}\left(J_H(N)^2\right)\\
\uparrow\wr\ v_2 & & & & \uparrow\wr\ v_1\\
\mathrm{Ta}_\mathfrak{m}\left(J_H(N)\right) & & & & \mathrm{Ta}_\mathfrak{m}\left(J_H(N)\right).
\end{array} \tag{2.12}$$

The vertical isomorphisms are defined by $v_2\colon x \to (-\langle p\rangle x, a_p x)$ and $v_1\colon x \to (a_p x,\ px)$. (Here $a_p \in \mathbf{T}_H(N)_\mathfrak{m}$ can be viewed as an element of $\mathbf{T}_H(N)_p \simeq \prod \mathbf{T}_H(N)_\mathfrak{n}$ where the product is taken over the maximal ideals containing p. So v_1 and v_2 can be viewed as maps to $\mathrm{Ta}_p\left(J_H(N)^2\right)$ whose images are respectively $\mathrm{Ta}_{\mathfrak{m}_1}\left(J_H(N)^2\right)$ and $\mathrm{Ta}_{\mathfrak{m}_2}\left(J_H(N)^2\right)$.)

Now $\widehat{\varphi}$ is surjective and φ is injective with torsion-free cokernel by the result of Ribet mentioned before. Also $\mathrm{Ta}_\mathfrak{m}\left(J_H(N)\right) \simeq \mathbf{T}_H(N)_\mathfrak{m}^2$ and

$\mathrm{Ta}_{\mathfrak{m}_p}\Big(J_H(N,p)\Big) \simeq \mathbf{T}_H(N,p)^2_{\mathfrak{m}_p}$ by Corollary 1 to Theorem 2.1. So as φ, $\widehat{\varphi}$ are maps of $\mathbf{T}_H(N,p)_{\mathfrak{m}_p}$-modules we can use this diagram to compute Δ_p as remarked just prior to the statement of the proposition. (The compatibility of the U_p actions requires that, on identifying the completions $S_{1,\mathfrak{m}_1}$ and $S_{2,\mathfrak{m}_2}$ with $\mathbf{T}_H(N)_\mathfrak{m}$, we get $U_1 = U_2$ which is indeed the case.) We find that

$$v_1^{-1} \circ \widehat{\varphi} \circ \varphi \circ v_2\,(z) \;=\; a_p^{-1}\,(a_p^2 - \langle p\rangle)\,(z). \qquad \square$$

We now apply to $J_1(N,q^2)$ (but $q \neq p$) the same analysis that we have just applied to $J_1(N,p)$. Here $X_1(A,B)$ is the curve corresponding to $\Gamma_1(A)\cap\Gamma_0(B)$ and $J_1(A,B)$ its Jacobian. First we need the analogue of Ihara's result. It is convenient to work in a slightly more general setting. Let us denote the maps $X_1(Nq^{r-1},q^r) \to X_1(Nq^{r-1})$ induced by $z \to z$ and $z \to qz$ by $\pi_{1,r}$ and $\pi_{2,r}$ respectively. Similarly we denote the maps $X_1(Nq^r,q^{r+1}) \to X_1(Nq^r)$ induced by $z \to z$ and $z \to qz$ by $\pi_{3,r}$ and $\pi_{4,r}$ respectively. Also let $\pi : X_1(Nq^r) \to X_1(Nq^{r-1},q^r)$ denote the natural map induced by $z \to z$.

In the following lemma if $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_1(Nq^{r-1})$ or $\mathbf{T}_1(Nq^r)$ we use $\mathfrak{m}^{(q)}$ to denote the maximal ideal of $\mathbf{T}_1^{(q)}(Nq^r,q^{r+1})$ compatible with $\mathfrak{m}$, the ring $\mathbf{T}_1^{(q)}(Nq^r,\,q^{r+1}) \subset \mathbf{T}_1(Nq^r,q^{r+1})$ being the subring obtained by omitting U_q from the list of generators.

LEMMA 2.5. *If $q \neq p$ is a prime and $r \geq 1$ then the sequence of abelian varieties*

$$0 \to J_1(Nq^{r-1}) \xrightarrow{\;\xi_1\;} J_1(Nq^r)\times J_1(Nq^r) \xrightarrow{\;\xi_2\;} J_1(Nq^r,\,q^{r+1})$$

where $\xi_1 = \Big((\pi_{1,r}\circ\pi)^*, -(\pi_{2,r}\circ\pi)^*\Big)$ and $\xi_2 = (\pi_{4,r}^*,\pi_{3,r}^*)$ induces a corresponding sequence of p-divisible groups which becomes exact when localized at any $\mathfrak{m}^{(q)}$ for which $\rho_\mathfrak{m}$ is irreducible.

Proof. Let $\Gamma^1(N\,q^r)$ denote the group $\Big\{\Big(\big(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\big)\Big) \in \Gamma_1(N) : a \equiv d \equiv 1(q^r),$ $c \equiv 0(q^{r-1}),\, b \equiv 0(q)\Big\}$. Let B_1 and B^1 be given by

$$B_1 = \Gamma_1(Nq^r)\big/\Gamma_1(Nq^r)\cap\Gamma(q), \qquad B^1 = \Gamma^1(Nq^r)\big/\Gamma_1(Nq^r)\cap\Gamma(q)$$

and let $\Delta_q = \Gamma_1(Nq^{r-1})\big/\Gamma_1(Nq^r)\cap\Gamma(q)$. Thus $\Delta_q \simeq \mathrm{SL}_2(\mathbf{Z}/q)$ if $r=1$ and is of order a power of q if $r>1$.

The exact sequences of inflation-restriction give:

$$H^1(\Gamma_1(N\,q^r),\mathbf{Q}_p/\mathbf{Z}_p) \xrightarrow[\sim]{\lambda_1} H^1(\Gamma_1(N\,q^r)\cap\Gamma(q),\mathbf{Q}_p/\mathbf{Z}_p)^{B_1},$$

together with a similar isomorphism with λ^1 replacing λ_1 and B^1 replacing B_1. We also obtain

$$H^1(\Gamma_1(Nq^{r-1}),\mathbf{Q}_p/\mathbf{Z}_p) \xrightarrow{\;\sim\;} H^1(\Gamma_1(Nq^r)\cap\Gamma(q),\mathbf{Q}_p/\mathbf{Z}_p)^{\Delta_q}.$$

The vanishing of $H^2(\mathrm{SL}_2(\mathbf{Z}/q), \mathbf{Q}_p/\mathbf{Z}_p)$ can be checked by restricting to the Sylow p-subgroup which is cyclic. Note that $\operatorname{im}\lambda_1 \cap \operatorname{im}\lambda^1 \subseteq H^1(\Gamma_1(N\,q^r) \cap \Gamma(q), \mathbf{Q}_p/\mathbf{Z}_p)^{\Delta_q}$ since B_1 and B^1 together generate Δ_q. Now consider the sequence

$$(2.13)\quad 0 \longrightarrow H^1(\Gamma_1(Nq^{r-1}), \mathbf{Q}_p/\mathbf{Z}_p) \xrightarrow{\mathrm{res}_1 \oplus -\mathrm{res}^1} H^1(\Gamma_1(Nq^r), \mathbf{Q}_p/\mathbf{Z}_p) \oplus H^1(\Gamma^1(Nq^r), \mathbf{Q}_p/\mathbf{Z}_p) \xrightarrow{\lambda_1 \oplus \lambda^1} H^1(\Gamma_1(Nq^r) \cap \Gamma(q), \mathbf{Q}_p/\mathbf{Z}_p).$$

We claim it is exact. To check this, suppose that $\lambda_1(x) = -\lambda^1(y)$. Then $\lambda_1(x) \in H^1(\Gamma_1(Nq^r) \cap \Gamma(q), \mathbf{Q}_p/\mathbf{Z}_p)^{\Delta_q}$. So $\lambda_1(x)$ is the restriction of an $x' \in H^1\Big(\Gamma_1(Nq^{r-1}), \mathbf{Q}_p/\mathbf{Z}_p\Big)$ whence $x - \mathrm{res}_1(x') \in \ker\lambda_1 = 0$. It follows also that $y = -\mathrm{res}^1(x')$.

Now conjugation by the matrix $\left(\begin{smallmatrix} q & 0 \\ 0 & 1 \end{smallmatrix}\right)$ induces isomorphisms

$$\Gamma^1(Nq^r) \simeq \Gamma_1(Nq^r), \qquad \Gamma_1(Nq^r) \cap \Gamma(q) \simeq \Gamma_1(Nq^r,\, q^{r+1}).$$

So our sequence (2.13) yields the exact sequence of the lemma, except that we have to change from group cohomology to the cohomology of the associated complete curves. If the groups are torsion-free then the difference between these cohomologies is Eisenstein (more precisely $T_l - 1 - l$ for $l \equiv 1 \bmod Nq^{r+1}$ is nilpotent) so will vanish when we localize at the preimage of $\mathfrak{m}^{(q)}$ in the abstract Hecke ring generated as a polynomial ring by all the standard Hecke operators excluding T_q. If $M \le 3$ then the group $\Gamma_1(M)$ has torsion. For $M = 1, 2, 3$ we can restrict to $\Gamma(3)$, $\Gamma(4)$, $\Gamma(3)$, respectively, where the cohomology is Eisenstein as the corresponding curves have genus zero and the groups are torsion-free. Thus one only needs to check the action of the Hecke operators on the kernels of the restriction maps in these three exceptional cases. This can be done explicitly and again they are Eisenstein. This completes the proof of the lemma. □

Let us denote the maps $X_1(N, q) \to X_1(N)$ induced by $z \to z$ and $z \to qz$ by π_1 and π_2 respectively. Similarly we denote the maps $X_1(N, q^2) \to X_1(N, q)$ induced by $z \to z$ and $z \to qz$ by π_3 and π_4 respectively.

From the lemma (with $r = 1$) and Ihara's result (2.10) we deduce that there is a sequence

$$(2.14)\qquad 0 \to J_1(N) \times J_1(N) \times J_1(N) \xrightarrow{\ \xi\ } J_1(N, q^2)$$

where $\xi = (\pi_1 \circ \pi_3)^* \times (\pi_2 \circ \pi_3)^* \times (\pi_2 \circ \pi_4)^*$ and that the induced map of p-divisible groups becomes injective after localization at $\mathfrak{m}^{(q)}$'s which correspond to irreducible $\rho_{\mathfrak{m}}$'s. By duality we obtain a sequence

$$J_1(N, q^2) \xrightarrow{\ \hat{\xi}\ } J_1(N)^3 \to 0$$

which is 'surjective' on Tate modules in the same sense. More generally we can prove analogous results for $J_H(N)$ and $J_H(N, q^2)$ although there may be

a kernel of order divisible by p in $J_H(N) \to J_1(N)$. However this kernel will not meet the $\mathfrak{m}^{(q)}$-divisible group for any maximal ideal $\mathfrak{m}^{(q)}$ whose associated $\rho_{\mathfrak{m}}$ is irreducible and hence, as in the earlier cases, will not affect the results if after passing to p-divisible groups we localize at such an $\mathfrak{m}^{(q)}$. We use the same notation in the general case when $H \neq 1$ so ξ is the map $J_H(N)^3 \to J_H(N, q^2)$.

We suppose now that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(N)$ (as always with $p \in \mathfrak{m}$) associated to an irreducible representation and that q is a prime, $q \nmid Np$. We now define a maximal ideal $\mathfrak{m}_q$ of $\mathbf{T}_H(N, q^2)$ with the same associated representation as $\mathfrak{m}$. To do this consider the ring

$$S_1 = \mathbf{T}_H(N)[U_1]/U_1(U_1^2 - T_q U_1 + q\langle q\rangle) \subseteq \operatorname{End}\left(J_H(N)^3\right)$$

where the action of U_1 on $J_H(N)^3$ is given by the matrix

$$\begin{bmatrix} T_q & -\langle q\rangle & 0 \\ q & 0 & 0 \\ 0 & q & 0 \end{bmatrix}.$$

Then U_1 satisfies the compatibility

$$\widehat{\xi} \circ U_q = U_1 \circ \widehat{\xi}$$

One checks this using the actions on cotangent spaces. For we may identify the cotangent spaces with spaces of cusp forms and with this identification any Hecke operator t_* induces the usual action on cusp forms. There is a maximal ideal $\mathfrak{m}_1 = (U_1, \mathfrak{m})$ in S_1 and $S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_{\mathfrak{m}}$. We let $\mathfrak{m}_q$ denote the reciprocal image of $\mathfrak{m}_1$ in $\mathbf{T}_H(N, q^2)$ under the natural map $\mathbf{T}_H(N, q^2) \to S_1$.

Next we define a principal ideal (Δ'_q) of $\mathbf{T}_H(N)_{\mathfrak{m}}$ using the fact that $\mathbf{T}_H(N, q^2)_{\mathfrak{m}_q}$ and $\mathbf{T}_H(N)_{\mathfrak{m}}$ are both Gorenstein rings (cf. Corollary 2 to Theorem 2.1). Thus we set $(\Delta'_q) = (\widehat{\alpha}' \circ \alpha')$ where

$$\alpha'\colon \mathbf{T}_H(N, q^2)_{\mathfrak{m}_q} \longrightarrow S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_{\mathfrak{m}}$$

is the natural map and $\widehat{\alpha}'$ is the adjoint with respect to selected Hecke-module pairings on $\mathbf{T}_H(N, q^2)_{\mathfrak{m}_q}$ and $\mathbf{T}_H(N)_{\mathfrak{m}}$. Note that α' is surjective. To show that the T_q operator is in the image one can use the existence of the associated 2-dimensional representation (cf. §1) in which $T_q = \operatorname{trace}(\operatorname{Frob} q)$ and apply the Čebotarev density theorem.

PROPOSITION 2.6. *Suppose that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(N)$ associated to an irreducible $\rho_{\mathfrak{m}}$. Suppose also that $q \nmid Np$. Then*

$$(\Delta'_q) = (q-1)\,(T_q^2 - \langle q\rangle(1+q)^2).$$

Proof. We prove this in the same manner as we proved Proposition 2.4. Consider the maps on p-adic Tate-modules induced by ξ and $\widehat{\xi}$:

$$\operatorname{Ta}_p\left(J_H(N)^3\right) \xrightarrow{\ \xi\ } \operatorname{Ta}_p\left(J_H(N, q^2)\right) \xrightarrow{\ \widehat{\xi}\ } \operatorname{Ta}_p\left(J_H(N)^3\right). \tag{2.15}$$

These maps commute with the standard Hecke operators with the exception of T_q and U_q (which are not even defined on all the terms). We define

$$S_2 = \mathbf{T}_H(N)[U_2]\Big/U_2(U_2^2 - T_q U_2 + q\langle q\rangle) \subseteq \operatorname{End}\left(J_H(N)^3\right)$$

where U_2 is the endomorphism of $J_H(N)^3$ given by the matrix

$$\begin{bmatrix} 0 & 0 & 0 \\ q & 0 & -\langle q\rangle \\ 0 & q & T_q \end{bmatrix}.$$

Then $U_q\,\xi = \xi\, U_2$ as one can verify by checking the equality $(\widehat{\xi}\circ\xi)U_2 = U_1(\widehat{\xi}\circ\xi)$ because $\widehat{\xi}\circ\xi$ is an isogeny. The formula for $\widehat{\xi}\circ\xi$ is given below. Again $\mathfrak{m}_2 = (\mathfrak{m}, U_2)$ is a maximal ideal of S_2 and $S_{2,\mathfrak{m}_2} \simeq \mathbf{T}_H(N)_\mathfrak{m}$. On restricting (2.15) to the $\mathfrak{m}_2$, $\mathfrak{m}_q$ and $\mathfrak{m}_1$-adic Tate modules we get

$$\begin{array}{ccccc} \mathrm{Ta}_{\mathfrak{m}_2}(J_H(N)^3) & \xrightarrow{\ \xi\ } & \mathrm{Ta}_{\mathfrak{m}_q}(J_H(N,\, q^2)) & \xrightarrow{\ \widehat{\xi}\ } & \mathrm{Ta}_{\mathfrak{m}_1}(J_H(N)^3) \\ \uparrow \wr\, u_2 & & & & \uparrow \wr\, u_1 \\ \mathrm{Ta}_\mathfrak{m}(J_H(N)) & & & & \mathrm{Ta}_\mathfrak{m}(J_H(N)). \end{array} \tag{2.16}$$

The vertical isomorphisms are induced by u_2: $z \longrightarrow (\langle q\rangle z, -T_q z, qz)$ and u_1: $z \longrightarrow (0,0,z)$. Now a calculation shows that on $J_H(N)^3$

$$\hat{\xi}\circ\xi = \begin{bmatrix} q(q+1) & T_q\cdot q & T_q^2 - \langle q\rangle(1+q) \\ T_q^*\cdot q & q(q+1) & T_q\cdot q \\ T_q^{*2} - \langle q\rangle^{-1}(1+q) & T_q^*\cdot q & q(q+1) \end{bmatrix}$$

where $T_q^* = \langle q\rangle^{-1}T_q$.

We compute then that

$$(u_1^{-1}\circ\widehat{\xi}\circ\xi\circ u_2) = -\langle q^{-1}\rangle(q-1)\left(T_q^2 - \langle q\rangle(1+q)^2\right).$$

Now using the surjectivity of $\widehat{\xi}$ and that ξ has torsion-free cokernel in (2.16) (by Lemma 2.5) and that $\mathrm{Ta}_\mathfrak{m}\left(J_H(N)\right)$ and $\mathrm{Ta}_{\mathfrak{m}_q}\left(J_H(N,\, q^2)\right)$ are each free of rank 2 over the respective Hecke rings (Corollary 1 of Theorem 2.1), we deduce the result as in Proposition 2.4. □

There is one further (and completely elementary) generalization of this result. We let π: $X_H(Nq,q^2) \to X_H(N,q^2)$ be the map given by $z \to z$. Then $\pi^*: J_H(N,q^2) \to J_H(Nq,q^2)$ has kernel a cyclic group and as before this will vanish when we localize at $\mathfrak{m}^{(q)}$ if $\mathfrak{m}$ is associated to an irreducible representation. (As before the superscript q denotes the omission of U_q from the list of generators of $\mathbf{T}_H(Nq,\, q^2)$ and $\mathfrak{m}^{(q)}$ denotes the maximal ideal of $\mathbf{T}_H^{(q)}(Nq,q^2)$ compatible with $\mathfrak{m}$.)

We thus have a sequence (not necessarily exact)

$$0 \to J_H(N)^3 \xrightarrow{\kappa} J_H(Nq,q^2) \to Z \to 0$$

where $\kappa = \pi^* \circ \xi$ which induces a corresponding sequence of p-divisible groups which becomes exact when localized at an $\mathfrak{m}^{(q)}$ corresponding to an irreducible $\rho_{\mathfrak{m}}$. Here Z is the quotient abelian variety $J_H(Nq, q^2)\big/\operatorname{im}\kappa$. As before there is a natural surjective homomorphism

$$\alpha\colon \mathbf{T}_H(Nq, q^2)_{\mathfrak{m}_q} \longrightarrow S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_{\mathfrak{m}}$$

where $\mathfrak{m}_q$ is the inverse image of $\mathfrak{m}_1$ in $\mathbf{T}_H(Nq, q^2)$. (We note that one can replace $\mathbf{T}_H(Nq,q^2)$ by $\mathbf{T}_H(Nq^2)$ in the definition of α and Proposition 2.7 below would still hold unchanged.) Since both rings are again Gorenstein we can define an adjoint $\widehat{\alpha}$ and a principal ideal

$$(\Delta_q) = (\alpha \circ \widehat{\alpha}) .$$

Proposition 2.7. *Suppose that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T} = \mathbf{T}_H(N)$ associated to an irreducible representation. Suppose that $q \nmid Np$. Then*

$$(\Delta_q) = (q-1)^2\Big(T_q^2 - \langle q\rangle(1+q)^2\Big).$$

The proof is a trivial generalization of that of Proposition 2.6.

Remark 2.8. We have included the operator U_q in the definition of $\mathbf{T}_{\mathfrak{m}_q} = \mathbf{T}_H(Nq, q^2)_{\mathfrak{m}_q}$ as in the application of the q-expansion principle it is important to have all the Hecke operators. However $U_q = 0$ in $\mathbf{T}_{\mathfrak{m}_q}$. To see this we recall that the absolute values of the eigenvalues $c(q,f)$ of U_q on newforms of level Nq with $q \nmid N$ are known (cf. [Li]). They satisfy $c(q,f)^2 = \langle q\rangle$ in $\mathcal{O}_f$ (the ring of integers generated by the Fourier coefficients of f) if f is on $\Gamma_1(N,q)$, and $|c(q,f)| = q^{1/2}$ if f is on $\Gamma_1(Nq)$ but not on $\Gamma_1(N,q)$. Also when f is a newform of level dividing N the roots of $x^2 - c(q,f)x + q\,\chi_f(q) = 0$ have absolute value $q^{1/2}$ where $c(q,f)$ is the eigenvalue of T_q and $\chi_f(q)$ of $\langle q\rangle$. Since for f on $\Gamma_1(Nq,q^2)$, $U_q f$ is a form on $\Gamma_1(Nq)$ we see that

$$U_q\,(U_q^2 - \langle q\rangle) \prod_{f\in\mathcal{S}_1} (U_q - c(q,f)) \prod_{f\in\mathcal{S}_2} \Big(U_q^2 - c(q,f)U_q + q\langle q\rangle\Big) = 0$$

in $\mathbf{T}_H(Nq, q^2) \otimes \mathbf{C}$ where $\mathcal{S}_1$ is the set of newforms on $\Gamma_1(Nq)$ which are not on $\Gamma_1(N,q)$ and $\mathcal{S}_2$ is the set of newforms of level dividing N. In particular as U_q is in $\mathfrak{m}_q$ it must be zero in $\mathbf{T}_{\mathfrak{m}_q}$.

A slightly different situation arises if $\mathfrak{m}$ is a maximal ideal of $\mathbf{T} = \mathbf{T}_H(N,q)$ $(q \neq p)$ which is not associated to any maximal ideal of level N (in the sense of having the same associated $\rho_{\mathfrak{m}}$). In this case we may use the map $\xi_3 = (\pi_4^*, \pi_3^*)$ to give

$$\text{(2.17)}\quad J_H(N,q) \times J_H(N,q) \xrightarrow{\xi_3} J_H(N,q^2) \xrightarrow{\hat{\xi}_3} J_H(N,q) \times J_H(N,q).$$

Then $\hat{\xi}_3 \circ \xi_3$ is given by the matrix

$$\hat{\xi}_3 \circ \xi_3 = \begin{bmatrix} q & U_q^* \\ & \\ U_q & q \end{bmatrix}$$

on $J_H(N,q)^2$, where $U_q^* = U_q\langle q\rangle^{-1}$ and $U_q^2 = \langle q\rangle$ on the $\mathfrak{m}$-divisible group. The second of these formulae is standard as mentioned above; cf. for example [Li, Th. 3], since $\rho_\mathfrak{m}$ is not associated to any maximal ideal of level N. For the first consider any newform f of level divisible by q and observe that the Petersson inner product $\left\langle (U_q^*U_q - 1)f(rz), f(mz)\right\rangle$ is zero for any $r, m \mid (Nq/\operatorname{level} f)$ by [Li, Th. 3]. This shows that $U_q^*U_q f(rz)$, *a priori* a linear combination of $f(m_i z)$, is equal to $f(rz)$. So $U_q^*U_q = 1$ on the space of forms on $\Gamma_H(N,q)$ which are new at q, i.e. the space spanned by forms $\{f(sz)\}$ where f runs through newforms with $q \mid \operatorname{level} f$. In particular U_q^* preserves the $\mathfrak{m}$-divisible group and satisfies the same relation on it, again because $\rho_\mathfrak{m}$ is not associated to any maximal ideal of level N.

Remark 2.9. Assume that $\rho_\mathfrak{m}$ is of type (A) at q in the terminology of Chapter 1, §1 (which ensures that $\rho_\mathfrak{m}$ does not occur at level N). In this case $\mathbf{T}_\mathfrak{m} = \mathbf{T}_H(N,\, q)_\mathfrak{m}$ is already generated by the standard Hecke operators with the omission of U_q. To see this, consider the $\mathrm{GL}_2\,(\mathbf{T}_\mathfrak{m})$ representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ associated to the $\mathfrak{m}$-adic Tate module of $J_H(N,q)$ (cf. the discussion following Corollary 2 of Theorem 2.1). Then this representation is already defined over the $\mathbf{Z}_p$-subalgebra $\mathbf{T}_\mathfrak{m}^{\mathrm{tr}}$ of $\mathbf{T}_\mathfrak{m}$ generated by the traces of Frobenius elements, i.e. by the T_ℓ for $\ell \nmid Nqp$. In particular $\langle q\rangle \in \mathbf{T}_\mathfrak{m}^{\mathrm{tr}}$. Furthermore, as $\mathbf{T}_\mathfrak{m}^{\mathrm{tr}}$ is local and complete, and as $U_q^2 = \langle q\rangle$, it is enough to solve $X^2 = \langle q\rangle$ in the residue field of $\mathbf{T}_\mathfrak{m}^{\mathrm{tr}}$. But we can even do this in k_0 (the minimal field of definition of $\rho_\mathfrak{m}$) by letting X be the eigenvalue of $\operatorname{Frob} q$ on the unique unramified rank-one free quotient of k_0^2 and invoking the $\pi_q \simeq \pi(\sigma_q)$ theorem of Langlands (cf. [Cal]). (It is to ensure that the unramified quotient is free of rank one that we assume $\rho_\mathfrak{m}$ to be of type (A).)

We assume now that $\rho_\mathfrak{m}$ is of type (A) at q. Define S_1 this time by setting

$$S_1 = \mathbf{T}_H(N,\, q)[U_1]\big/U_1(U_1 - U_q) \subseteq \operatorname{End}\left(J_H(N,\, q)^2\right)$$

where U_1 is given by the matrix

$$U_1 = \begin{bmatrix} 0 & q \\ 0 & U_q \end{bmatrix} \tag{2.18}$$

on $J_H(N,\, q)^2$. The map $\widehat{\xi}_3$ is not necessarily surjective and to remedy this we introduce $\mathfrak{m}^{(q)} = \mathfrak{m} \cap \mathbf{T}_H^{(q)}(N,q)$ where $\mathbf{T}_H^{(q)}(N,q)$ is the subring of $\mathbf{T}_H(N,q)$ generated by the standard Hecke operators but omitting U_q. We also write $\mathfrak{m}^{(q)}$

for the corresponding maximal ideal of $\mathbf{T}_H^{(q)}(Nq, q^2)$. Then on $\mathfrak{m}^{(q)}$-divisible groups, $\widehat{\xi}_3$ and $\widehat{\xi}_3 \circ \pi_*$ are surjective and we get a natural restriction map of localizations $\mathbf{T}_H(Nq, q^2)_{(\mathfrak{m}^{(q)})} \to S_{1(\mathfrak{m}^{(q)})}$. (Note that the image of U_q under this map is U_1 and not U_q.) The ideal $\mathfrak{m}_1 = (\mathfrak{m}, U_1)$ is maximal in S_1 and so also in $S_{1,(\mathfrak{m}^{(q)})}$ and we let $\mathfrak{m}_q$ denote the inverse image of $\mathfrak{m}_1$ under this restriction map. The inverse image of $\mathfrak{m}_q$ in $\mathbf{T}_H(Nq, q^2)$ is also a maximal ideal which we again write $\mathfrak{m}_q$. Since the completions $\mathbf{T}_H(Nq,\, q^2)_{\mathfrak{m}_q}$ and $S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N,\, q)_{\mathfrak{m}}$ are both Gorenstein rings (by Corollary 2 of Theorem 2.1) we can define a principal ideal (Δ_q) of $\mathbf{T}_H(N,\, q)_{\mathfrak{m}}$ by

$$(\Delta_q) = (\alpha \circ \widehat{\alpha})$$

where α: $\mathbf{T}_H(Nq,\, q^2)_{\mathfrak{m}_q} \twoheadrightarrow S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N,\, q)_{\mathfrak{m}}$ is the restriction map induced by the restriction map on $\mathfrak{m}^{(q)}$-localizations described above.

PROPOSITION 2.10. *Suppose that* $\mathfrak{m}$ *is a maximal ideal of* $\mathbf{T}_H(N, q)$ *associated to an irreducible* $\mathfrak{m}$ *of type (A). Then*

$$(\Delta_q) = (q-1)^2\,(q+1).$$

Proof. The method is a straightforward adaptation of that used for Propositions 2.4 and 2.6. We let $S_2 = \mathbf{T}_H(N,\, q)[U_2]\big/U_2(U_2 - U_q)$ be the ring of endomorphisms of $J_H(N, q)^2$ where U_2 is given by the matrix

$$\begin{bmatrix} U_q & q \\ 0 & 0 \end{bmatrix}.$$

This satisfies the compatibility $\xi_3 U_2 = U_q\, \xi_3$. We define $\mathfrak{m}_2 = (\mathfrak{m}, U_2)$ in S_2 and observe that $S_2, \mathfrak{m}_2 \simeq \mathbf{T}_H(N, q)_{\mathfrak{m}}$.

Then we have maps

$$\begin{array}{ccccc} \mathrm{Ta}_{\mathfrak{m}_2}\Big(J_H(N,\, q)^2\Big) & \overset{\pi^* \circ \xi_3}{\hookrightarrow} & \mathrm{Ta}_{\mathfrak{m}_q}\Big(J_H(Nq,\, q^2)\Big) & \overset{\hat{\xi}_3 \circ \pi_*}{\twoheadrightarrow} & \mathrm{Ta}_{\mathfrak{m}_1}\Big(J_H(N,\, q)^2\Big) \\ \uparrow\wr \quad v_2 & & & & \uparrow\wr \quad v_1 \\ \mathrm{Ta}_{\mathfrak{m}}\Big(J_H(N,\, q)\Big) & & & & \mathrm{Ta}_{\mathfrak{m}}\Big(J_H(N,\, q)\Big). \end{array}$$

The maps v_1 and v_2 are given by v_2: $z \to (-qz,\ a_q z)$ and v_1: $z \to (z,\, 0)$ where $U_q = a_q$ in $\mathbf{T}_H(N,\, q)_{\mathfrak{m}}$. One checks then that $v_1^{-1} \circ (\hat{\xi}_3 \circ \pi_*) \circ (\pi^* \circ \xi_3) \circ v_2$ is equal to $-(q-1)\,(q^2-1)$ or $-\frac{1}{2}(q-1)(q^2-1)$.

The surjectivity of $\widehat{\xi}_3 \circ \pi_*$ on the completions is equivalent to the statement that

$$J_H(Nq, q^2)[p]_{\mathfrak{m}_q} \longrightarrow J_H(N, q)^2[p]_{\mathfrak{m}_1}$$

is surjective. We can replace this condition by a similar one with $\mathfrak{m}^{(q)}$ substituted for $\mathfrak{m}_q$ and for $\mathfrak{m}_1$, i.e., the surjectivity of

$$J_H(Nq, q^2)[p]_{\mathfrak{m}^{(q)}} \longrightarrow J_H(N, q)^2[p]_{\mathfrak{m}^{(q)}}.$$

By our hypothesis that $\rho_{\mathfrak{m}}$ be of type (A) at q it is even sufficient to show that the cokernel of $J_H(Nq, q^2)[p] \otimes \overline{\mathbf{F}}_p \to J_H(N, q)^2[p] \otimes \overline{\mathbf{F}}_p$ has no subquotient as a Galois-module which is irreducible, two-dimensional and ramified at q. This statement, or rather its dual, follows from Lemma 2.5. The injectivity of $\pi^* \circ \xi_3$ on the completions and the fact that it has torsion-free cokernel also follows from Lemma 2.5 and our hypothesis that $\rho_{\mathfrak{m}}$ be of type (A) at q. □

The case that corresponds to type (B) is similar. We assume in the analysis of type (B) (and also of type (C) below) that H decomposes as $\prod H_q$ as described at the beginning of Section 1. We assume that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(Nq^r)$ where H contains the Sylow p-subgroup S_p of $(\mathbf{Z}/q^r\mathbf{Z})^*$ and that

$$\rho_{\mathfrak{m}}\Big|_{I_q} \approx \begin{pmatrix} \chi_q & \\ & 1 \end{pmatrix} \tag{2.19}$$

for a suitable choice of basis with $\chi_q \neq 1$ and $\operatorname{cond}\chi_q = q^r$. Here $q \nmid Np$ and we assume also that $\rho_{\mathfrak{m}}$ is irreducible. We use the sequence

$$J_H(Nq^r) \times J_H(Nq^r) \xrightarrow{(\pi')^* \circ \xi_2} J_{H'}(Nq^r, q^{r+1}) \xrightarrow{\hat{\xi}_2 \circ \pi'_*} J_H(Nq^r) \times J_H(Nq^r)$$

defined analogously to (2.17) where ξ_2 was as defined in Lemma 2.5 and where H' is defined as follows. Using the notation $H = \prod H_l$ as at the beginning of Section 1 set $H'_l = H_l$ for $l \neq q$ and $H'_q \times S_p = H_q$. Then define $H' = \prod H'_l$ and let $\pi' : X_{H'}(Nq^r,\, q^{r+1}) \to X_H(Nq^r,\, q^{r+1})$ be the natural map $z \to z$. Using Lemma 2.5 we check that ξ_2 is injective on the $\mathfrak{m}^{(q)}$-divisible group. Again we set $S_1 = \mathbf{T}_H(Nq^r)\,[U_1]\big/U_1\,(U_1 - U_q) \subseteq \operatorname{End}(J_H\big(Nq^r)^2\big)$ where U_1 is given by the matrix in (2.18). We define $\mathfrak{m}_1 = (\mathfrak{m},\, U_1)$ and let $\mathfrak{m}_q$ be the inverse image of $\mathfrak{m}_1$ in $\mathbf{T}_{H'}(Nq^r,\, q^{r+1})$. The natural map (in which $U_q \to U_1$)

$$\alpha\colon\ \mathbf{T}_{H'}(Nq^r,\, q^{r+1})_{\mathfrak{m}_q} \longrightarrow S_{1,\,\mathfrak{m}_1} \simeq \mathbf{T}_H(Nq^r)_{\mathfrak{m}}$$

is surjective by the following remark.

Remark 2.11. When we assume that $\rho_{\mathfrak{m}}$ is of type (B) then the U_q operator is redundant in $\mathbf{T}_{\mathfrak{m}} = \mathbf{T}_H(Nq^r)_{\mathfrak{m}}$. To see this, first assume that $\mathbf{T}_{\mathfrak{m}}$ is reduced and consider the $\mathrm{GL}_2(\mathbf{T}_{\mathfrak{m}})$ representation of $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ associated to the $\mathfrak{m}$-adic Tate module. Pick a $\sigma_q \in I_q$, the inertia group in D_q in $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, such that $\chi_q(\sigma_q) \neq 1$. Then because the eigenvalues of σ_q are distinct $\operatorname{mod} \mathfrak{m}$ we can diagonalize the representation with respect to σ_q. If $\operatorname{Frob} q$ is a Frobenius in D_q, then in the $\mathrm{GL}_2(\mathbf{T}_{\mathfrak{m}})$ representation the image of $\operatorname{Frob} q$ normalizes I_q and we can recover U_q as the entry of the matrix giving the value of $\operatorname{Frob} q$ on the unit eigenvector for σ_q. This is by the $\pi_q \simeq \pi(\sigma_q)$ theorem of Langlands as before (cf. [Cal]) applied to each of the representations obtained from maps $\mathbf{T}_{\mathfrak{m}} \to \mathcal{O}_{f,\lambda}$. Since the representation is defined over the $\mathbf{Z}_p$-algebra $\mathbf{T}_{\mathfrak{m}}^{\mathrm{tr}}$ generated by the traces, the same reasoning applied to $\mathbf{T}_{\mathfrak{m}}^{\mathrm{tr}}$ shows that $U_q \in \mathbf{T}_{\mathfrak{m}}^{\mathrm{tr}}$.

If $\mathbf{T}_\mathfrak{m}$ is not reduced the above argument shows only that there is an operator $v_q \in \mathbf{T}_\mathfrak{m}^{\text{tr}}$ such that $(U_q - v_q)$ is nilpotent. Now $\mathbf{T}_H(Nq^r)$ can be viewed as a ring of endomorphisms of $S_2(\Gamma_H(Nq^r))$, the space of cusp forms of weight 2 on $\Gamma_H(Nq^r)$. There is a restriction map $\mathbf{T}_H(Nq^r) \to \mathbf{T}_H(Nq^r)^{\text{new}}$ where $\mathbf{T}_H(Nq^r)^{\text{new}}$ is the image of $\mathbf{T}_H(Nq^r)$ in the ring of endomorphisms of $S_2(\Gamma_H(Nq^r))/S_2(\Gamma_H(Nq^r))^{\text{old}}$, the old part being defined as the sum of two copies of $S_2(\Gamma_H(Nq^{r-1}))$ mapped via $z \to z$ and $z \to qz$. One sees that on $\mathfrak{m}$-completions $\mathbf{T}_\mathfrak{m} \simeq (\mathbf{T}_H(Nq^r)^{\text{new}})_\mathfrak{m}$ since the conductor of $\rho_\mathfrak{m}$ is divisible by q^r. It follows that $U_q \in \mathbf{T}_\mathfrak{m}$ satisfies an equation of the form $P(U_q) = 0$ where $P(x)$ is a polynomial with coefficients in $W(k_\mathfrak{m})$ and with distinct roots. By extending scalars to $\mathcal{O}$ (the integers of a local field containing $W(k_\mathfrak{m})$) we can assume that the roots lie in $T \simeq \mathbf{T}_\mathfrak{m} \underset{W(k_\mathfrak{m})}{\otimes} \mathcal{O}$.

Since $(U_q - v_q)$ is nilpotent it follows that $P(v_q)^r = 0$ for some r. Then since $v_q \in \mathbf{T}_\mathfrak{m}^{\text{tr}}$ which is reduced, $P(v_q) = 0$. Now consider the map $T \to \Pi T_{(\mathfrak{p})}$ where the product is taken over the localizations of T at the minimal primes $\mathfrak{p}$ of T. The map is injective since the associated primes of the kernel are all maximal, whence the kernel is of finite cardinality and hence zero. Now in each $T_{(\mathfrak{p})}$, $U_q = \alpha_i$ and $v_q = \alpha_j$ for roots α_i, α_j of $P(x) = 0$ because the roots are distinct. Since $U_q - v_q \in \mathfrak{p}$ for each $\mathfrak{p}$ it follows that $\alpha_i = \alpha_j$ for each $\mathfrak{p}$ whence $U_q = v_q$ in each $T_{(\mathfrak{p})}$. Hence $U_q = v_q$ in T also and this finally shows that $U_q \in \mathbf{T}_\mathfrak{m}^{\text{tr}}$ in general.

We can therefore define a principal ideal

$$(\Delta_q) = (\alpha \circ \widehat{\alpha})$$

using, as previously, that the rings $\mathbf{T}_{H'}(Nq^r, q^{r+1})_{\mathfrak{m}_q}$ and $\mathbf{T}_H(Nq^r)_\mathfrak{m}$ are Gorenstein. We compute (Δ_q) in a similar manner to the type (A) case, but using this time that $U_q^* U_q = q$ on the space of forms on $\Gamma_H(Nq^r)$ which are new at q, i.e., the space spanned by forms $\{f(sz)\}$ where f runs through newforms with $q^r \mid \text{level} f$. To see this let f be any newform of level divisible by q^r and observe that the Petersson inner product $\Big\langle (U_q^* U_q - q) f(rz), f(mz) \Big\rangle = 0$ for any $m \mid (Nq^r/\text{level} f)$ by [Li, Th. 3(ii)]. This shows that $(U_q^* U_q - q) f(rz)$, *a priori* a linear combination of $\{f(m_i z)\}$, is zero. We obtain the following result.

PROPOSITION 2.12. *Suppose that* $\mathfrak{m}$ *is a maximal ideal of* $\mathbf{T}_H(Nq^r)$ *associated to an irreducible* $\rho_\mathfrak{m}$ *of type* (B) *at* q, *i.e., satisfying* (2.19) *including the hypothesis that* H *contains* S_p. *(Again* $q \nmid Np$.*) Then*

$$(\Delta_q) = \Big((q-1)^2\Big).$$

Finally we have the case where $\rho_\mathfrak{m}$ is of type (C) at q. We assume then that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(Nq^r)$ where H contains the Sylow p-subgroup

S_p of $(\mathbf{Z}/q^r\mathbf{Z})^*$ and that

$$H^1(\mathbf{Q}_q, W_\lambda) = 0 \tag{2.20}$$

where W_λ is defined as in (1.6) but with $\rho_\mathfrak{m}$ replacing ρ_0, i.e., $W_\lambda = \mathrm{ad}^0\,\rho_\mathfrak{m}$.

This time we let $\mathfrak{m}_q$ be the inverse image of $\mathfrak{m}$ in $\mathbf{T}_{H'}(Nq^r)$ under the natural restriction map $\mathbf{T}_{H'}(Nq^r) \longrightarrow \mathbf{T}_H(Nq^r)$ with H' defined as in the case of type B. We set

$$(\Delta_q) = (\alpha \circ \hat{\alpha})$$

where α: $\mathbf{T}_{H'}(Nq^r)_{\mathfrak{m}_q} \twoheadrightarrow \mathbf{T}_H(Nq^r)_\mathfrak{m}$ is the induced map on the completions, which as before are Gorenstein rings. The proof of the following proposition is analogous (but simpler) to the proof of Proposition 2.10. (Notice that the proposition does not require the condition that $\rho_\mathfrak{m}$ satisfy (2.20) but this is the case in which we will use it.)

PROPOSITION 2.13. *Suppose that $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(Nq^r)$ associated to an irreducible $\rho_\mathfrak{m}$ with H containing the Sylow p-subgroup of $(\mathbf{Z}/q^r\mathbf{Z})^*$. Then*

$$(\Delta_q) = (q-1).$$

Finally, in this section we state Proposition 2.4 in the case $q \neq p$ as this will be used in Chapter 3. Let q be a prime, $q \nmid Np$ and let S_1 denote the ring

$$T_H(N)[U_1]/\{U_1^2 - T_qU_1 + \langle q\rangle q\} \subseteq \mathrm{End}(J_H(N)^2) \tag{2.21}$$

where $\hat{\varphi}\colon J_H(N,q) \to J_H(N)^2$ is the map defined after (2.10) and U_1 is the matrix

$$\begin{bmatrix} T_q & -\langle q\rangle \\ q & 0 \end{bmatrix}.$$

Thus, $\hat{\varphi}U_q = U_1\hat{\varphi}$. Also $\langle q\rangle$ is defined as $\langle n_q\rangle$ where $n_q \equiv q(N)$, $n_q \equiv 1(q)$. Let $\mathfrak{m}_1$ be a maximal ideal of S_1 containing the image of $\mathfrak{m}$, where $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(N)$ with associated irreducible $\rho_\mathfrak{m}$. We will also assume that $\rho_\mathfrak{m}(\mathrm{Frob}\,q)$ has distinct eigenvalues. (We will only need this case and it simplifies the exposition.) Let $\mathfrak{m}_q$ denote the corresponding maximal ideals of $\mathbf{T}_H(N,q)$ and $\mathbf{T}_H(Nq)$ under the natural restriction maps $\mathbf{T}_H(Nq) \to \mathbf{T}_H(N,q) \to S_1$. The corresponding maps on completions are

$$\begin{aligned} \mathbf{T}_H(Nq)_{\mathfrak{m}_q} &\xrightarrow{\beta} \mathbf{T}_H(N,q)_{\mathfrak{m}_q} \\ &\xrightarrow{\alpha} S_{1,\mathfrak{m}_1} \simeq \mathbf{T}_H(N)_\mathfrak{m} \underset{W(k_\mathfrak{m})}{\otimes} W(k^+) \end{aligned} \tag{2.22}$$

where k^+ is the extension of $k_\mathfrak{m}$ generated by the eigenvalues of $\{\rho_\mathfrak{m}(\mathrm{Frob}\,q)\}$. Thus k^+ is either equal to $k_\mathfrak{m}$ or its quadratic extension. The maps β, α are surjective, the latter because T_q is a trace in the 2-dimensional representation

to $\mathrm{GL}_2(\mathbf{T}_H(N)_{\mathfrak{m}})$ given after Theorem 2.1 and hence is 'redundant' by the Čebotarev density theorem. The completions are Gorenstein by Corollary 2 to Theorem 2.1 and so we define invariant ideals of $S_{1,\mathfrak{m}_1}$

$$(\Delta) = (\alpha \circ \hat{\alpha}), \qquad (\Delta') = (\alpha \circ \beta) \circ (\widehat{\alpha \circ \beta}). \tag{2.23}$$

Let α_q be the image of U_1 in $\mathbf{T}_H(N)_{\mathfrak{m}} \underset{W(k_{\mathfrak{m}})}{\otimes} W(k^+)$ under the last isomorphism in (2.22). The proof of Proposition 2.4 yields

PROPOSITION 2.4′. *Suppose that $\rho_{\mathfrak{m}}$ is irreducible where $\mathfrak{m}$ is a maximal ideal of $\mathbf{T}_H(N)$ and that $\rho_{\mathfrak{m}}(\mathrm{Frob}\, q)$ has distinct eigenvalues. Then*

$$\begin{aligned} (\Delta) &= (\alpha_q^2 - \langle q \rangle), \\ (\Delta') &= (\alpha_q^2 - \langle q \rangle)(q-1). \end{aligned}$$

Remark. Note that if we suppose also that $q \equiv 1(p)$ then (Δ) is the unit ideal and α is an isomorphism in (2.22).

3. The main conjectures

As we suggested in Chapter 1, in order to study the deformation theory of ρ_0 in detail we need to assume that it is modular. That this should always be so for $\det \rho_0$ odd was conjectured by Serre. Serre also made a conjecture (the 'ε'-conjecture) making precise where one could find a lifting of ρ_0 once one assumed it to be modular (cf. [Se]). This has now been proved by the combined efforts of a number of authors including Ribet, Mazur, Carayol, Edixhoven and others. The most difficult step was to show that if ρ_0 was unramified at a prime l then one could find a lifting in which l did not divide the level. This was proved (in slightly less generality) by Ribet. For a precise statement and complete references we refer to Diamond's paper [Dia] which removed the last restrictions referred to in Ribet's survey article [Ri3]. The following is a minor adaptation of the epsilon conjecture to our situation which can be found in [Dia, Th. 6.4]. (We wish to use weight 2 only.) Let $N(\rho_0)$ be the prime to p part of the conductor of ρ_0 as defined for example in [Se].

THEOREM 2.14. *Suppose that ρ_0 is modular and satisfies* (1.1) (*so in particular is irreducible*) *and is of type $\mathcal{D} = (\cdot,\Sigma,\mathcal{O},\mathcal{M})$ with* $\cdot =$ Se, str *or* fl. *Suppose that at least one of the following conditions holds* (i) $p > 3$ *or* (ii) ρ_0 *is not induced from a character of $\mathbf{Q}(\sqrt{-3})$. Then there exists a newform f of weight 2 and a prime λ of $\mathcal{O}_f$ such that $\rho_{f,\lambda}$ is of type $\mathcal{D}' = (\cdot,\Sigma,\mathcal{O}',\mathcal{M})$ for some $\mathcal{O}'$, and such that $(\rho_{f,\lambda} \bmod \lambda) \simeq \rho_0$ over $\overline{\mathbf{F}}_p$. Moreover we can assume that f has character χ_f of order prime to p and has level $N(\rho_0)p^{\delta(\rho_0)}$*

where $\delta(\rho_0) = 0$ *if* $\rho_0|_{D_p}$ *is associated to a finite flat group scheme over* $\mathbf{Z}_p$ *and* $\det \rho_0\big|_{I_p} = \omega$, *and* $\delta(\rho_0) = 1$ *otherwise. Furthermore in the Selmer case we can assume that* $a_p(f) \equiv \chi_2(\mathrm{Frob}\, p) \bmod \lambda$ *in the notation of* (1.2) *where* $a_p(f)$ *is the eigenvalue of* U_p.

For the rest of this chapter we will assume that ρ_0 is modular and that if $p = 3$ then ρ_0 is not induced from a character of $\mathbf{Q}(\sqrt{-3})$. Here and in the rest of the paper we use the term 'induced' to signify that the representation is induced after an extension of scalars to the algebraic closure.

For each $\mathcal{D} = \{\cdot, \Sigma, \mathcal{O}, \mathcal{M}\}$ we will now define a Hecke ring $\mathbf{T}_\mathcal{D}$ except where $\cdot$ is unrestricted. Suppose first that we are in the flat, Selmer or strict cases. Recall that when referring to the flat case we assume that ρ_0 is not ordinary and that $\det \rho_0|_{I_p} = \omega$. Suppose that $\Sigma = \{q_i\}$ and that $N(\rho_0) = \Pi\, q_i^{s_i}$ with $s_i \geq 0$. If $U_\lambda \simeq k^2$ is the representation space of ρ_0 we set $n_q = \dim_k (U_\lambda)^{I_q}$ where I_q is the inertia group at q. Define M_0 and M by

$$(2.24) \qquad M_0 = N(\rho_0) \prod_{\substack{n_{q_i}=1 \\ q_i \notin \mathcal{M}\cup\{p\}}} q_i \cdot \prod_{n_{q_i}=2} q_i^2, \qquad M = M_0\, p^{\tau(\rho_0)}$$

where $\tau(\rho_0) = 1$ if ρ_0 is ordinary and $\tau(\rho_0) = 0$ otherwise. Let H be the subgroup of $(\mathbf{Z}/M\mathbf{Z})^*$ generated by the Sylow p-subgroup of $(\mathbf{Z}/q_i\mathbf{Z})^*$ for each $q_i \in \mathcal{M}$ as well as by all of $(\mathbf{Z}/q_i\mathbf{Z})^*$ for each $q_i \in \mathcal{M}$ of type (A). Let $\mathbf{T}'_H(M)$ denote the ring generated by the standard Hecke operators $\{T_l$ for $l \nmid Mp$, $\langle a \rangle$ for $(a,\ Mp) = 1\}$. Let $\mathfrak{m}'$ denote the maximal ideal of $\mathbf{T}'_H(M)$ associated to the f and λ given in the theorem and let $k_{\mathfrak{m}'}$ be the residue field $\mathbf{T}'_H(M)/\mathfrak{m}'$. Note that $\mathfrak{m}'$ does not depend on the particular choice of pair (f, λ) in theorem 2.14. Then $k_{\mathfrak{m}'} \simeq k_0$ where k_0 is the smallest possible field of definition for ρ_0 because $k_{\mathfrak{m}'}$ is generated by the traces. Henceforth we will identify k_0 with $k_{\mathfrak{m}'}$. There is one exceptional case where ρ_0 is ordinary and $\rho_0|_{D_p}$ is isomorphic to a sum of two distinct unramified characters (χ_1 and χ_2 in the notation of Chapter 1, §1). If ρ_0 is not exceptional we define

$$(2.25(\mathrm{a})) \qquad \mathbf{T}_\mathcal{D} = \mathbf{T}'_H(M)_{\mathfrak{m}'} \underset{W(k_0)}{\otimes} \mathcal{O}.$$

If ρ_0 is exceptional we let $\mathbf{T}''_H(M)$ denote the ring generated by the operators $\{T_l$ for $l \nmid Mp$, $\langle a \rangle$ for $(a,\ Mp) = 1,\ U_p\}$. We choose $\mathfrak{m}''$ to be a maximal ideal of $\mathbf{T}''_H(M)$ lying above $\mathfrak{m}'$ for which there is an embedding $k_{\mathfrak{m}''} \hookrightarrow k$ (over $k_0 = k_{\mathfrak{m}'}$) satisfying $U_p \to \chi_2(\mathrm{Frob}\, p)$. (Note that χ_2 is specified by $\mathcal{D}$.) Then in the exceptional case $k_{\mathfrak{m}''}$ is either k_0 or its quadratic extension and we define

$$(2.25(\mathrm{b})) \qquad \mathbf{T}_\mathcal{D} = \mathbf{T}''_H(M)_{\mathfrak{m}''} \underset{W(k_{\mathfrak{m}''})}{\otimes} \mathcal{O}.$$

The omission of the Hecke operators U_q for $q \mid M_0$ ensures that $\mathbf{T}_\mathcal{D}$ is reduced.

We need to relate $\mathbf{T}_{\mathcal{D}}$ to a Hecke ring with no missing operators in order to apply the results of Section 1.

PROPOSITION 2.15. *In the nonexceptional case there is a maximal ideal* $\mathfrak{m}$ *for* $\mathbf{T}_H(M)$ *with* $\mathfrak{m} \cap \mathbf{T}'_H(M) = \mathfrak{m}'$ *and* $k_0 = k_{\mathfrak{m}}$, *and such that the natural map* $\mathbf{T}'_H(M)_{\mathfrak{m}'} \to \mathbf{T}_H(M)_{\mathfrak{m}}$ *is an isomorphism, thus giving*

$$\mathbf{T}_{\mathcal{D}} \simeq \mathbf{T}_H(M)_{\mathfrak{m}} \underset{W(k_0)}{\otimes} \mathcal{O}.$$

In the exceptional case the same statements hold with $\mathfrak{m}''$ *replacing* $\mathfrak{m}'$, $\mathbf{T}''_H(M)$ *replacing* $\mathbf{T}'_H(M)$ *and* $k_{\mathfrak{m}''}$ *replacing* k_0.

Proof. For simplicity we describe the nonexceptional case indicating where appropriate the slight modifications needed in the exceptional case. To construct $\mathfrak{m}$ we take the eigenform f_0 obtained from the newform f of Theorem 2.14 by removing the Euler factors at all primes $q \in \Sigma - \{\mathcal{M} \cup p\}$. If ρ_0 is ordinary and f has level prime to p we also remove the Euler factor $(1 - \beta_p \cdot p^{-s})$ where β_p is the non-unit eigenvalue in $\mathcal{O}_{f,\lambda}$. (By 'removing Euler factors' we mean take the eigenform whose L-series is that of f with these Euler factors removed.) Then f_0 is an eigenform of weight 2 on $\Gamma_H(M)$ (this is ensured by the choice of f) with $\mathcal{O}_{f,\lambda}$ coefficients. We have a corresponding homomorphism $\pi_{f_0}\colon \mathbf{T}_H(M) \to \mathcal{O}_{f,\lambda}$ and we let $\mathfrak{m} = \pi_{f_0}^{-1}(\lambda)$.

Since the Hecke operators we have used to generate $\mathbf{T}'_H(M)$ are prime to the level there is an inclusion with finite index

$$\mathbf{T}'_H(M) \hookrightarrow \prod \mathcal{O}_g$$

where g runs over representatives of the Galois conjugacy classes of newforms associated to $\Gamma_H(M)$ and where we note that by multiplicity one $\mathcal{O}_g$ can also be described as the ring of integers generated by the eigenvalues of the operators in $\mathbf{T}'_H(M)$ acting on g. If we consider $\mathbf{T}_H(M)$ in place of $\mathbf{T}'_H(M)$ we get a similar map but we have to replace the ring $\mathcal{O}_g$ by the ring

$$S_g = \mathcal{O}_g[X_{q_1}, \ldots, X_{q_r}, X_p] \big/ \{Y_i, Z_p\}_{i=1}^r$$

where $\{p, q_1, \ldots, q_r\}$ are the distinct primes dividing Mp. Here

$$\text{(2.26)} \qquad Y_i = \begin{cases} X_{q_i}^{r_i - 1}\left(X_{q_i} - \alpha_{q_i}(g)\right)\left(X_{q_i} - \beta_{q_i}(g)\right) & \text{if } q_i \nmid \text{level}(g) \\ X_{q_i}^{r_i}\left(X_{q_i} - a_{q_i}(g)\right) & \text{if } q_i \mid \text{level}(g), \end{cases}$$

where the Euler factor of g at q_i (i.e., of its associated L-series) is $(1 - \alpha_{q_i}(g) q_i^{-s})(1 - \beta_{q_i}(g) q_i^{-s})$ in the first case and $(1 - a_{q_i}(g) q_i^{-s})$ in the second case, and $q_i^{r_i} \| \left(M/\text{level}(g)\right)$. (We allow $a_{q_i}(g)$ to be zero here.) Similarly Z_p is

defined by

$$Z_p = \begin{cases} X_p^2 - a_p(g)X_p + p\chi_g(p) & \text{if } p \mid M,\ p \nmid \operatorname{level}(g) \\ X_p - a_p(g) & \text{if } p \nmid M \\ X_p - a_p(g) & \text{if } p \mid \operatorname{level}(g), \end{cases}$$

where the Euler factor of g at p is $(1 - a_p(g)p^{-s} + \chi_g(p)\,p^{1-2s})$ in the first two cases and $(1 - a_p(g)p^{-s})$ in the third case. We then have a commutative diagram

$$(2.27)\qquad \begin{array}{ccc} \mathbf{T}'_H(M) & \hookrightarrow & \prod_g \mathcal{O}_g \\ \big\downarrow & & \big\downarrow \\ \mathbf{T}_H(M) & \hookrightarrow & \prod_g S_g = \prod_g \mathcal{O}_g[X_{q_1}, \ldots, X_{q_r}, X_p]/\{Y_i, Z_p\}_{i=1}^r \end{array}$$

where the lower map is given on $\{U_{q_i},\ U_p \text{ or } T_p\}$ by $U_{q_i} \longrightarrow X_{q_i}$, U_p or $T_p \longrightarrow X_p$ (according as $p \mid M$ or $p \nmid M$). To verify the existence of such a homomorphism one considers the action of $\mathbf{T}_H(M)$ on the space of forms of weight 2 invariant under $\Gamma_H(M)$ and uses that $\sum_{j=1}^r g_j(m_j z)$ is a free generator as a $\mathbf{T}_H(M) \otimes \mathbf{C}$-module where $\{g_j\}$ runs over the set of newforms and $m_j = M/\operatorname{level}(g_j)$.

Now we tensor all the rings in (2.27) with $\mathbf{Z}_p$. Then completing the top row of (2.27) with respect to $\mathfrak{m}'$ and the bottom row with respect to $\mathfrak{m}$ we get a commutative diagram

$$(2.28)\qquad \begin{array}{ccccc} \mathbf{T}'_H(M)_{\mathfrak{m}'} & \hookrightarrow & \Big(\prod_g \mathcal{O}_g\Big)_{\mathfrak{m}'} & \simeq & \prod_{\substack{g \\ \mathfrak{m}' \to \mu}} \mathcal{O}_{g,\mu} \\ \big\downarrow & & \big\downarrow & & \big\downarrow \\ \mathbf{T}_H(M)_{\mathfrak{m}} & \hookrightarrow & \Big(\prod_g S_g\Big)_{\mathfrak{m}} & \simeq & \prod (S_g)_{\mathfrak{m}}. \end{array}$$

Here μ runs through the primes above p in each $\mathcal{O}_g$ for which $\mathfrak{m}' \to \mu$ under $\mathbf{T}_{H'}(M) \to \mathcal{O}_g$. Now $(S_g)_{\mathfrak{m}}$ is given by

$$\begin{aligned} (2.29)\quad (S_g \otimes \mathbf{Z}_p)_{\mathfrak{m}} &\simeq \Big((\mathcal{O}_g \otimes \mathbf{Z}_p)\,[X_{q_1}, \ldots, X_{q_r},\ X_p]\big/\{Y_i,\ Z_p\}_{i=1}^r\Big)_{\mathfrak{m}} \\ &\simeq \Bigg(\prod_{\mu \mid p} \mathcal{O}_{g,\mu}\,[X_{q_1}, \ldots, X_{q_r},\ X_p]\big/\{Y_i,\ Z_p\}_{i=1}^r\Bigg)_{\mathfrak{m}} \\ &\simeq \Bigg(\prod_{\mu \mid p} A_{g,\mu}\Bigg)_{\mathfrak{m}} \end{aligned}$$

where $A_{g,\mu}$ denotes the product of the factors of the complete semi-local ring $\mathcal{O}_{g,\mu}[X_{q_1}, \ldots, X_{q_r}, X_p]\big/\{Y_i,\ Z_p\}_{i=1}^r$ in which X_{q_i} is topologically nilpotent for

$q_i \notin \mathcal{M}$ and in which X_p is a unit if we are in the ordinary case (i.e., when $p \mid M$). This is because $U_{q_i} \in \mathfrak{m}$ if $q_i \notin \mathcal{M}$ and U_p is a unit at $\mathfrak{m}$ in the ordinary case.

Now if $\mathfrak{m}' \to \mu$ then in $(A_{g,\mu})_\mathfrak{m}$ we claim that Y_i is given up to a unit by $X_{q_i} - b_i$ for some $b_i \in \mathcal{O}_{g,\mu}$ with $b_i = 0$ if $q_i \notin \mathcal{M}$. Similarly Z_p is given up to a unit by $X_p - \alpha_p(g)$ where $\alpha_p(g)$ is the unit root of $x^2 - a_p(g)x + p\chi_g(p) = 0$ in $\mathcal{O}_{g,\mu}$ if $p \nmid$ level g and $p \mid M$ and by $X_p - a_p(g)$ if $p \mid$ level g or $p \nmid M$. This will show that $(A_{g,\mu})_\mathfrak{m} \simeq \mathcal{O}_{g,\mu}$ when $\mathfrak{m}' \to \mu$ and $(A_{g,\mu})_\mathfrak{m} = 0$ otherwise.

For $q_i \in \mathcal{M}$ and for p, the claim is straightforward. For $q_i \notin \mathcal{M}$, it amounts to the following. Let $U_{g,\mu}$ denote the 2-dimensional $K_{g,\mu}$-vector space with Galois action via $\rho_{g,\mu}$ and let $n_{q_i}(g,\mu) = \dim(U_{g,\mu})^{I_{q_i}}$. We wish to check that $Y_i =$ unit. X_{q_i} in $(A_{g,\mu})_\mathfrak{m}$ and from the definition of Y_i in (2.26) this reduces to checking that $r_i = n_{q_i}(g,\mu)$ by the $\pi_q \simeq \pi(\sigma_q)$ theorem (cf. [Ca1]). We use here that $\alpha_{q_i}(g)$, $\beta_{q_i}(g)$ and $a_{q_i}(g)$ are p-adic units when they are nonzero since they are eigenvalues of $\operatorname{Frob} q_i$. Now by definition the power of q_i dividing M is the same as that dividing $N(\rho_0)q_i^{n_{q_i}}$ (cf. (2.21)). By an observation of Livné (cf. [Liv], [Ca2, §1]),

$$\operatorname{ord}_{q_i}(\operatorname{level} g) = \operatorname{ord}_{q_i}\left(N(\rho_0)q_i^{n_{q_i} - n_{q_i}(g,\mu)}\right). \tag{2.30}$$

As by definition $q_i^{r_i} \| (M/\operatorname{level} g)$ we deduce that $r_i = n_{q_i}(g,\mu)$ as required.

We have now shown that each $A_{g,\mu} \simeq \mathcal{O}_{g,\mu}$ (when $\mathfrak{m}' \to \mu$) and it follows from (2.28) and (2.29) that we have homomorphisms

$$\mathbf{T}'_H(M)_{\mathfrak{m}'} \hookrightarrow \mathbf{T}_H(M)_\mathfrak{m} \hookrightarrow \prod_{\substack{g \\ \mathfrak{m}' \to \mu}} \mathcal{O}_{g,\mu}$$

where the inclusions are of finite index. Moreover we have seen that $U_{q_i} = 0$ in $\mathbf{T}_H(M)_\mathfrak{m}$ for $q_i \notin \mathcal{M}$. We now consider the primes $q_i \in \mathcal{M}$. We have to show that the operators U_q for $q \in \mathcal{M}$ are redundant in the sense that they lie in $\mathbf{T}'_H(M)_{\mathfrak{m}'}$, i.e., in the $\mathbf{Z}_p$-subalgebra of $\mathbf{T}_H(M)_\mathfrak{m}$ generated by the $\{T_l\colon l \nmid Mp,\ \langle a \rangle\colon a \in (\mathbf{Z}/M\mathbf{Z})^*\}$. For $q \in \mathcal{M}$ of type (A), $U_q \in \mathbf{T}'_H(M)_{\mathfrak{m}'}$ as explained in Remark 2.9 and for $q \in \mathcal{M}$ of type (B), $U_q \in \mathbf{T}'_H(M)_{\mathfrak{m}'}$ as explained in Remark 2.11. For $q \in \mathcal{M}$ of type (C) but not of type (A), $U_q = 0$ by the $\pi_q \simeq \pi(\sigma_q)$ theorem (cf. [Ca1]). For in this case $n_q = 0$ whence also $n_q(g,\mu) = 0$ for each pair (g,μ) with $\mathfrak{m}' \to \mu$. If ρ_0 is strict or Selmer at p then U_p can be recovered from the two-dimensional representation ρ (described after the corollaries to Theorem 2.1) as the eigenvalue of $\operatorname{Frob} p$ on the (free, of rank one) unramified quotient (cf. Theorem 2.1.4 of [Wil]). As this representation is defined over the $\mathbf{Z}_p$-subalgebra generated by the traces, it follows that U_p is contained in this subring. In the exceptional case U_p is in $\mathbf{T}''_H(M)_{\mathfrak{m}''}$ by definition.

Finally we have to show that T_p is also redundant in the sense explained above when $p \nmid M$. A proof of this has already been given in Section 2 (Ribet's

lemma). Here we give an alternative argument using the Galois representations. We know that $T_p \in \mathfrak{m}$ and it will be enough to show that $T_p \in (\mathfrak{m}^2, p)$. Writing $k_{\mathfrak{m}}$ for the residue field $\mathbf{T}_H(M)_{\mathfrak{m}}/\mathfrak{m}$ we reduce to the following situation. If $T_p \notin (\mathfrak{m}^2, p)$ then there is a quotient

$$\mathbf{T}_H(M)_{\mathfrak{m}}/(\mathfrak{m}^2, p) \twoheadrightarrow k_{\mathfrak{m}}[\varepsilon] = \mathbf{T}_H(M)_{\mathfrak{m}}/\mathfrak{a}$$

where $k_{\mathfrak{m}}[\varepsilon]$ is the ring of dual numbers (so $\varepsilon^2 = 0$) with the property that $T_p \mapsto \lambda\varepsilon$ with $\lambda \neq 0$ and such that the image of $\mathbf{T}'_H(M)_{\mathfrak{m}'}$ lies in $k_{\mathfrak{m}}$. Let $G_{/\mathbf{Q}}$ denote the four-dimensional $k_{\mathfrak{m}}$-vector space associated to the representation

$$\rho_\varepsilon\colon \operatorname{Gal}(\overline{\mathbf{Q}}/\mathbf{Q}) \longrightarrow \operatorname{GL}_2(k_{\mathfrak{m}}[\varepsilon])$$

induced from the representation in Theorem 2.1. It has the form

$$G_{/\mathbf{Q}} \simeq G_{0/\mathbf{Q}} \oplus G_{0/\mathbf{Q}}$$

where G_0 is the corresponding space associated to ρ_0 by our hypothesis that the traces lie in $k_{\mathfrak{m}}$. The semisimplicity of $G_{/\mathbf{Q}}$ here is obtained from the main theorem of [BLR]. Now $G_{/\mathbf{Q}_p}$ extends to a finite flat group scheme $G_{/\mathbf{Z}_p}$. Explicitly it is a quotient of the group scheme $J_H(M)_{\mathfrak{m}}[p]_{/\mathbf{Z}_p}$. Since extensions to $\mathbf{Z}_p$ are unique (cf. [Ray1]) we know

$$G_{/\mathbf{Z}_p} \simeq G_{0/\mathbf{Z}_p} \oplus G_{0/\mathbf{Z}_p}.$$

Now by the Eichler-Shimura relation we know that in $J_H(M)_{/\mathbf{F}_p}$

$$T_p = F + \langle p\rangle F^T.$$

Since $T_p \in \mathfrak{m}$ it follows that $F + \langle p\rangle F^T = 0$ on $G_{0/\mathbf{F}_p}$ and hence the same holds on $G_{/\mathbf{F}_p}$. But T_p is an endomorphism of $G_{/\mathbf{Z}_p}$ which is zero on the special fibre, so by [Ray1, Cor. 3.3.6], $T_p = 0$ on $G_{/\mathbf{Z}_p}$. It follows that $T_p = 0$ in $k_{\mathfrak{m}}[\varepsilon]$ which contradicts our earlier hypothesis. So $T_p \in (\mathfrak{m}^2, p)$ as required. This completes the proof of the proposition. □

From the proof of the proposition it is also clear that $\mathfrak{m}$ is the unique maximal ideal of $\mathbf{T}_H(M)$ extending $\mathfrak{m}'$ and satisfying the conditions that $U_q \in \mathfrak{m}$ for $q \in \Sigma - \{\mathcal{M} \cup p\}$ and $U_p \notin \mathfrak{m}$ if ρ_0 is ordinary. For the rest of this chapter we will always make this choice of $\mathfrak{m}$ (given ρ_0).

Next we define $\mathbf{T}_{\mathcal{D}}$ in the case when $\mathcal{D} = (\text{ord}, \Sigma, \mathcal{O}, \mathcal{M})$. If $\mathfrak{n}$ is any ordinary maximal ideal (i.e. $U_p \notin \mathfrak{n}$) of $\mathbf{T}_H(Np)$ with N prime to p then Hida has constructed a 2-dimensional Noetherian local Hecke ring

$$\mathbf{T}_\infty = e\ \mathbf{T}_H(Np^\infty)_{\mathfrak{n}} := \varprojlim\ e\ \mathbf{T}_H(Np^r)_{\mathfrak{n}_r}$$

which is a $\Lambda = \mathbf{Z}_p[\![T]\!]$-algebra satisfying $\mathbf{T}_\infty/T \simeq \mathbf{T}_H(Np)_{\mathfrak{n}}$. Here $\mathfrak{n}_r$ is the inverse image of $\mathfrak{n}$ under the natural restriction map. Also $T = \varprojlim \langle 1 + Np\rangle - 1$

and $e = \varinjlim_{r} U_p^{r!}$. For an irreducible ρ_0 of type $\mathcal{D}$ we have defined $\mathbf{T}_{\mathcal{D}'}$ in (2.25(a)), where $\mathcal{D}' = (\mathrm{Se}, \Sigma, \mathcal{O}, \mathcal{M})$, by

$$\mathbf{T}_{\mathcal{D}'} \simeq \mathbf{T}_H(M_0 p)_{\mathfrak{m}} \underset{W(k_{\mathfrak{m}})}{\otimes} \mathcal{O},$$

the isomorphism coming from Proposition 2.15. We will define $\mathbf{T}_{\mathcal{D}}$ by

$$\mathbf{T}_{\mathcal{D}} = e\mathbf{T}_H(M_0 p^\infty)_{\mathfrak{m}} \underset{W(k_{\mathfrak{m}})}{\otimes} \mathcal{O}. \tag{2.31}$$

In particular we see that

$$\mathbf{T}_{\mathcal{D}}/T \simeq \mathbf{T}_{\mathcal{D}'}, \tag{2.32}$$

i.e., where $\mathcal{D}'$ is the same as $\mathcal{D}$ but with 'Selmer' replacing 'ord'. Moreover if $\mathfrak{q}$ is a height one prime ideal of $\mathbf{T}_{\mathcal{D}}$ containing $\left((1+T)^{p^n} - (1+Np)^{p^n(k-2)}\right)$ for any integers $n \geq 0$, $k \geq 2$, then $\mathbf{T}_{\mathcal{D}}/\mathfrak{q}$ is associated to an eigenform in a natural way (generalizing the case $n = 0$, $k = 2$). For more details about these rings as well as about Λ-adic modular forms see for example [Wi1] or [Hi1].

For each $n \geq 1$ let $\mathbf{T}_n = \mathbf{T}_H(M_0 p^n)_{\mathfrak{m}_n}$. Then by the argument given after the statement of Theorem 2.1 we can construct a Galois representation ρ_n unramified outside Mp with values in $\mathrm{GL}_2(\mathbf{T}_n)$ satisfying trace $\rho_n(\mathrm{Frob}\, l) = T_l$, $\det \rho_n(\mathrm{Frob}\, l) = l\langle l \rangle$ for $(l, Mp) = 1$. These representations can be patched together to give a continuous representation

$$\rho = \varprojlim \rho_n \colon \mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathbf{T}_{\mathcal{D}}) \tag{2.33}$$

where Σ is the set of primes dividing Mp. To see this we need to check the commutativity of the maps

$$\begin{array}{ccc} R_\Sigma & \longrightarrow & \mathbf{T}_n \\ & \searrow & \downarrow \\ & & \mathbf{T}_{n-1} \end{array}$$

where the horizontal maps are induced by ρ_n and ρ_{n-1} and the vertical map is the natural one. Now the commutativity is valid on elements of R_Σ, which are traces or determinants in the universal representation, since trace $(\mathrm{Frob}\, l) \mapsto T_l$ under both horizontal maps and similarly for determinants. Here R_Σ is the universal deformation ring described in Chapter 1 with respect to ρ_0 viewed with residue field $k = k_{\mathfrak{m}}$. It suffices then to show that R_Σ is generated (topologically) by traces and this reduces to checking that there are no nonconstant deformations of ρ_0 to $k[\varepsilon]$ with traces lying in k (cf. [Ma1, §1.8]). For then if R_Σ^{tr} denotes the closed $W(k)$-subalgebra of R_Σ generated by the traces we see that $R_\Sigma^{\mathrm{tr}} \to (R_\Sigma/m^2)$ is surjective, m being the maximal ideal of R_Σ, from which we easily conclude that $R_\Sigma^{\mathrm{tr}} = R_\Sigma$. To see that the condition holds, assume

that a basis is chosen so that $\rho_0(c) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ for a chosen complex conjugation c and $\rho_0(\sigma) = \begin{pmatrix} a_\sigma & b_\sigma \\ c_\sigma & d_\sigma \end{pmatrix}$ with $b_\sigma = 1$ and $c_\sigma \neq 0$ for some σ. (This is possible because ρ_0 is irreducible.) Then any deformation $[\rho]$ to $k[\varepsilon]$ can be represented by a representation ρ such that $\rho(c)$ and $\rho(\sigma)$ have the same properties. It follows easily that if the traces of ρ lie in k then ρ takes values in k whence it is equal to ρ_0. (Alternatively one sees that the universal representation can be defined over R_Σ^{tr} by diagonalizing complex conjugation as before. Since the two maps $R_\Sigma^{\mathrm{tr}} \to \mathbf{T}_{n-1}$ induced by the triangle are the same, so the associated representations are equivalent, and the universal property then implies the commutativity of the triangle.)

The representations (2.33) were first exhibited by Hida and were the original inspiration for Mazur's deformation theory.

For each $\mathcal{D} = \{\cdot, \Sigma, \mathcal{O}, \mathcal{M}\}$ where $\cdot$ is not unrestricted there is then a canonical surjective map

$$\varphi_{\mathcal{D}} : R_{\mathcal{D}} \to \mathbf{T}_{\mathcal{D}}$$

which induces the representations described after the corollaries to Theorem 2.1 and in (2.33). It is enough to check this when $\mathcal{O} = W(k_0)$ (or $W(k_{\mathfrak{m}''})$ in the exceptional case). Then one just has to check that for every pair (g, μ) which appears in (2.28) the resulting representation is of type $\mathcal{D}$. For then we claim that the image of the canonical map $R_{\mathcal{D}} \longrightarrow \widetilde{\mathbf{T}_{\mathcal{D}}} = \Pi\, \mathcal{O}_{g,\mu}$ is $\mathbf{T}_{\mathcal{D}}$ where here $\sim$ denotes the normalization. (In the case where $\cdot$ is ord this needs to be checked instead for $\mathbf{T}_n \underset{W(k_0)}{\otimes} \mathcal{O}$ for each n.) For this we just need to see that $R_{\mathcal{D}}$ is generated by traces. (In the exceptional case we have to show also that U_p is in the image. This holds because it can be identified, using Theorem 2.1.4 of [Wi1], with the image of $u \in R_{\mathcal{D}}$ where u is the eigenvalue of $\mathrm{Frob}\, p$ on the unique rank one unramified quotient of $R_{\mathcal{D}}^2$ with eigenvalue $\equiv \chi_2(\mathrm{Frob}\, p)$ which is specified in the definition of $\mathcal{D}$.) But we saw above that this was true for R_Σ. The same then holds for $R_{\mathcal{D}}$ as $R_\Sigma \longrightarrow R_{\mathcal{D}}$ is surjective because the map on reduced cotangent spaces is surjective (cf. (1.5)). To check the condition on the pairs (g, μ) observe first that for $q \in \mathcal{M}$ we have imposed the following conditions on the level and character of such g's by our choice of M and H:

q of type (A): $q \| \operatorname{level} g$, $\det \rho_{g,\mu}\big|_{I_q} = 1$,

q of type (B): $\operatorname{cond} \chi_q \| \operatorname{level} g$, $\det \rho_{g,\mu}\big|_{I_q} = \chi_q$,

q of type (C): $\det \rho_{g,\mu}\big|_{I_q}$ is the Teichmüller lifting of $\det \rho_0\big|_{I_q}$.

In the first two cases the desired form of $\rho_{g,\mu}\big|_{D_q}$ then follows from the $\pi_q \simeq \pi(\sigma_q)$ theorem of Langlands (cf. [Ca1]). The third case is already of

type (C). For $q = p$ one can use Theorem 2.1.4 of [Wil] in the ordinary case, the flat case being well-known.

The following conjecture generalizes a fundamental conjecture of Mazur and Tilouine for $\mathcal{D} = (\text{ord}, \Sigma, W(k_0), \phi)$; cf. [MT].

CONJECTURE 2.16. *$\varphi_{\mathcal{D}}$ is an isomorphism.*

Equivalently this conjecture says that the representation described after the corollaries to Theorem 2.1 (or in (2.33) in the ordinary case) is the universal one for a suitable choice of H, N and $\mathfrak{m}$. We remind the reader that throughout this section we are assuming that if $p = 3$ then ρ_0 is not induced from a character of $\mathbf{Q}(\sqrt{-3})$.

Remark. The case of most interest to us is when $p = 3$ and ρ_0 is a representation with values in $\mathrm{GL}_2(\mathbf{F}_3)$. In this case it is a theorem of Tunnell, extending results of Langlands, that ρ_0 is always modular. For $\mathrm{GL}_2(\mathbf{F}_3)$ is a double cover of S_4 and can be embedded in $\mathrm{GL}_2(\mathbf{Z}[\sqrt{-2}])$ whence in $\mathrm{GL}_2(\mathbf{C})$; cf. [Se] and [Tu]. The conjecture will be proved with a mild restriction on ρ_0 at the end of Chapter 3.

Remark. Our original restriction to the types (A), (B), (C) for ρ_0 was motivated by the wish that the deformation type (a) be of minimal conductor among its twists, (b) retain property (a) under unramified base changes. Without this kind of stability it can happen that after a base change of $\mathbf{Q}$ to an extension unramified at Σ, $\rho_0 \otimes \psi$ has smaller 'conductor' for some character ψ. The typical example of this is where $\rho_0\big|_{D_q} = \mathrm{Ind}_K^{\mathbf{Q}_q}(\chi)$ with $q \equiv -1(p)$ and χ is a ramified character over K, the unramified quadratic extension of $\mathbf{Q}_q$. What makes this difficult for us is that there are then nontrivial ramified local deformations ($\mathrm{Ind}_K^{\mathbf{Q}_p} \chi\xi$ for ξ a ramified character of order p of K) which we cannot detect by a change of level.

For the purposes of Chapter 3 it is convenient to digress now in order to introduce a slight variant of the deformation rings we have been considering so far. Suppose that $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}, \mathcal{M})$ is a standard deformation problem (associated to ρ_0) with $\cdot =$ Se, str or fl and suppose that H, M_0, M and $\mathfrak{m}$ are defined as in (2.24) and Proposition 2.15. We choose a finite set of primes $Q = \{q_1, \dots, q_r\}$ with $q_i \nmid Mp$. Furthermore we assume that each $q_i \equiv 1(p)$ and that the eigenvalues $\{\alpha_i, \beta_i\}$ of $\rho_0(\mathrm{Frob}\, q_i)$ are distinct for each $q_i \in Q$. This last condition ensures that ρ_0 does not occur as the residual representation of the λ-adic representation associated to any newform on $\Gamma_H(M, q_1 \cdot \ldots \cdot q_r)$ where any q_i divides the level of the form. This can be seen directly by looking at $(\mathrm{Frob}\, q_i)$ in such a representation or by using Proposition 2.4$'$ at the end of Section 2. It will be convenient to assume that the residue field of $\mathcal{O}$ contains α_i, β_i for each q_i.

Pick α_i for each i. We let $\mathcal{D}_Q$ be the deformation problem associated to representations ρ of $\mathrm{Gal}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q})$ which are of type $\mathcal{D}$ and which in addition satisfy the property that at each $q_i \in Q$

$$\rho|_{D_{q_i}} \sim \begin{pmatrix} \chi_{1,q_i} & \\ & \chi_{2,q_i} \end{pmatrix} \tag{2.34}$$

with χ_{2,q_i} unramified and $\chi_{2,q_i}(\mathrm{Frob}\, q_i) \equiv \alpha_i \bmod \mathfrak{m}$ for a suitable choice of basis. One checks as in Chapter 1 that associated to $\mathcal{D}_Q$ there is a universal deformation ring R_Q. (These new conditions are really variants on type (B).)

We will only need a corresponding Hecke ring in a very special case and it is convenient in this case to define it using all the Hecke operators. Let us now set $N = N(\rho_0)p^{\delta(\rho_0)}$ where $\delta(\rho_0)$ is as defined in Theorem 2.14. Let $\mathfrak{m}_0$ denote a maximal ideal of $\mathbf{T}_H(N)$ given by Theorem 2.14 with the property that $\rho_{\mathfrak{m}_0} \simeq \rho_0$ over $\overline{\mathbf{F}}_p$ relative to a suitable embedding of $k_{\mathfrak{m}_0} \to k$ over k_0. (In the exceptional case we also impose the same condition on $\mathfrak{m}_0$ about the reduction of U_p as in the definition of $\mathbf{T}_{\mathcal{D}}$ in the exceptional case before (2.25)(b).) Thus $\rho_{\mathfrak{m}_0} \simeq \rho_{f,\lambda} \bmod \lambda$ over the residue field of $\mathcal{O}_{f,\lambda}$ for some choice of f and λ with f of level N. By dropping one of the Euler factors at each q_i as in the proof of Proposition 2.15, we obtain a form and hence a maximal ideal $\mathfrak{m}_Q$ of $\mathbf{T}_H(Nq_1 \dots q_r)$ with the property that $\rho_{\mathfrak{m}_Q} \simeq \rho_0$ over $\overline{\mathbf{F}}_p$ relative to a suitable embedding $k_{\mathfrak{m}_Q} \to k$ over $k_{\mathfrak{m}_0}$. The field $k_{\mathfrak{m}_Q}$ is the extension of k_0 (or $k_{\mathfrak{m}''}$ in the exceptional case) generated by the α_i, β_i. We set

$$\mathbf{T}_Q = \mathbf{T}_H(Nq_1 \dots q_r)_{\mathfrak{m}_Q} \underset{W(k_{\mathfrak{m}_Q})}{\otimes} \mathcal{O}. \tag{2.35}$$

It is easy to see directly (or by the arguments of Proposition 2.15) that $\mathbf{T}_Q$ is reduced and that there is an inclusion with finite index

$$\mathbf{T}_Q \hookrightarrow \widetilde{\mathbf{T}}_Q = \prod \mathcal{O}_{g,\mu} \tag{2.36}$$

where the product is taken over representatives of the Galois conjugacy classes of eigenforms g of level $Nq_1 \dots q_r$ with $\mathfrak{m}_Q \to \mu$. Now define $\mathcal{D}_Q$ using the choices α_i for which $U_{q_i} \to \alpha_i$ under the chosen embedding $k_{\mathfrak{m}_Q} \to k$. Then each of the 2-dimensional representations associated to each factor $\mathcal{O}_{g,\mu}$ is of type $\mathcal{D}_Q$. We can check this for each $q \in Q$ using either the $\pi_q \simeq \pi(\sigma_q)$ theorem (cf. [Ca1]) as in the case of type (B) or using the Eichler-Shimura relation if q does not divide the level of the newform associated to g. So we get a homomorphism of $\mathcal{O}$-algebras $R_Q \to \widetilde{\mathbf{T}}_Q$ and hence also an $\mathcal{O}$-algebra map

$$\varphi_Q \colon R_Q \to \mathbf{T}_Q \tag{2.37}$$

as R_Q is generated by traces. This is not an isomorphism in general as we have used N in place of M. However it is surjective by the arguments of Proposition 2.15. Indeed, for $q \mid N(\rho_0)p$, we check that U_q is in the image of

φ_Q using the arguments in the second half of the proof of Proposition 2.15. For $q \in Q$ we use the fact that U_q is the image of the value of $\chi_{2,q}(\operatorname{Frob} q)$ in the universal representation; cf. (2.34). For $q \mid M$, but not of the previous types, T_q is a trace in $\rho_{\mathbf{T}_Q}$ and we can apply the Čebotarev density theorem to show that it is in the image of φ_Q.

Finally, if there is a section $\pi\colon \mathbf{T}_Q \to \mathcal{O}$, then set $\mathfrak{p}_Q = \ker \pi$ and let $\rho_{\mathfrak{p}}$ denote the 2-dimensional representation to $\mathrm{GL}_2(\mathcal{O})$ obtained from $\rho_{\mathbf{T}_Q} \bmod \mathfrak{p}_Q$. Let $V = \operatorname{Ad}\rho_{\mathfrak{p}} \otimes_{\mathcal{O}} K/\mathcal{O}$ where K is the field of fractions of $\mathcal{O}$. We pick a basis for $\rho_{\mathfrak{p}}$ satisfying (2.34) and then let

$$
\begin{aligned}
(2.38) \qquad V^{(q_i)} &= \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \right\} \\
\subseteq \operatorname{Ad}\rho_{\mathfrak{p}} \otimes_{\mathcal{O}} K/\mathcal{O} &= \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a,b,c,d \in \mathcal{O} \right\} \otimes_{\mathcal{O}} K/\mathcal{O}
\end{aligned}
$$

and let $V_{(q_i)} = V/V^{(q_i)}$. Then as in Proposition 1.2 we have an isomorphism

$$
(2.39) \qquad \operatorname{Hom}_{\mathcal{O}}(\mathfrak{p}_{R_Q}/\mathfrak{p}_{R_Q}^2, K/\mathcal{O}) \simeq H^1_{\mathcal{D}_Q}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V)
$$

where $\mathfrak{p}_{R_Q} = \ker(\pi \circ \varphi_Q)$ and the second term is defined by

$$
(2.40) \qquad H^1_{\mathcal{D}_Q}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V) = \ker\colon H^1_{\mathcal{D}}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V) \to \prod_{i=1}^{r} H^1(\mathbf{Q}_{q_i}^{\mathrm{unr}}, V_{(q_i)}).
$$

We return now to our discussion of Conjecture 2.16. We will call a deformation theory $\mathcal{D}$ *minimal* if $\Sigma = \mathcal{M} \cup \{p\}$ and $\cdot$ is Selmer, strict or flat. This notion will be critical in Chapter 3. (A slightly stronger notion of minimality is described in Chapter 3 where the Selmer condition is replaced, when possible, by the condition that the representations arise from finite flat group schemes—see the remark after the proof of Theorem 3.1.) Unfortunately even up to twist, not every ρ_0 has an associated minimal $\mathcal{D}$ even when ρ_0 is flat or ordinary at p as explained in the remarks after Conjecture 2.16. However this could be achieved if one replaced $\mathbf{Q}$ by a suitable finite extension depending on ρ_0.

Suppose now that f is a (normalized) newform, λ is a prime of $\mathcal{O}_f$ above p and $\rho_{f,\lambda}$ a deformation of ρ_0 of type $\mathcal{D}$ where $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}_{f,\lambda}, \mathcal{M})$ with $\cdot = \mathrm{Se}$, str or fl. (Strictly speaking we may be changing ρ_0 as we wish to choose its field of definition to be $k = \mathcal{O}_{f,\lambda}/\lambda$.) Suppose further that level$(f) \mid M$ where M is defined by (2.24).

Now let us set $\mathcal{O} = \mathcal{O}_{f,\lambda}$ for the rest of this section. There is a homomorphism

$$
(2.41) \qquad \pi = \pi_{\mathcal{D},f} : \mathbf{T}_{\mathcal{D}} \to \mathcal{O}
$$

whose kernel is the prime ideal $\mathfrak{p}_{\mathbf{T},f}$ associated to f and λ. Similarly there is a homomorphism

$$R_{\mathcal{D}} \to \mathcal{O}$$

whose kernel is the prime ideal $\mathfrak{p}_{R,f}$ associated to f and λ and which factors through π_f. Pick perfect pairings of $\mathcal{O}$-modules, the second one $\mathbf{T}_{\mathcal{D}}$-bilinear,

$$\mathcal{O} \times \mathcal{O} \to \mathcal{O}, \qquad \langle\,,\rangle : \mathbf{T}_{\mathcal{D}} \times \mathbf{T}_{\mathcal{D}} \to \mathcal{O}. \tag{2.42}$$

In each case we use the term perfect pairing to signify that the pairs of induced maps $\mathcal{O} \to \mathrm{Hom}_{\mathcal{O}}(\mathcal{O}, \mathcal{O})$ and $\mathbf{T}_{\mathcal{D}} \to \mathrm{Hom}_{\mathcal{O}}(\mathbf{T}_{\mathcal{D}}, \mathcal{O})$ are isomorphisms. In addition the second one is required to be $\mathbf{T}_{\mathcal{D}}$-linear. The existence of the second pairing is equivalent to the Gorenstein property, Corollary 2 of Theorem 2.1, as we explain below. Explicitly if h is a generator of the free $\mathbf{T}_{\mathcal{D}}$-module $\mathrm{Hom}_{\mathcal{O}}(\mathbf{T}_{\mathcal{D}}, \mathcal{O})$ we set $\langle t_1, t_2\rangle = h(t_1 t_2)$.

A priori $\mathbf{T}_H(M)_{\mathfrak{m}}$ (occurring in the description of $\mathbf{T}_{\mathcal{D}}$ in Proposition 2.15) is only Gorenstein as a $\mathbf{Z}_p$-algebra but it follows immediately that it is also a Gorenstein $W(k_{\mathfrak{m}})$-algebra. (The notion of Gorenstein $\mathcal{O}$-algebra is explained in the appendix.) Indeed the map

$$\mathrm{Hom}_{W(k_{\mathfrak{m}})}\left(\mathbf{T}_H(M)_{\mathfrak{m}},\ W(k_{\mathfrak{m}})\right) \longrightarrow \mathrm{Hom}_{\mathbf{Z}_p}\left(\mathbf{T}_H(M)_{\mathfrak{m}},\ \mathbf{Z}_p\right)$$

given by $\varphi \mapsto \mathrm{trace} \circ \varphi$ is easily seen to be an isomorphism, as the reduction $\bmod p$ is injective and the ranks are equal. Thus $\mathbf{T}_{\mathcal{D}}$ is a Gorenstein $\mathcal{O}$-algebra.

Now let $\hat{\pi} : \mathcal{O} \to \mathbf{T}_{\mathcal{D}}$ be the adjoint of π with respect to these pairings. Then define a principal ideal (η) of $\mathbf{T}_{\mathcal{D}}$ by

$$(\eta) = (\eta_{\mathcal{D},f}) = (\hat{\pi}(1)).$$

This is well-defined independently of the pairings and moreover one sees that $\mathbf{T}_{\mathcal{D}}/\eta$ is torsion-free (see the appendix). From its description (η) is invariant under extensions of $\mathcal{O}$ to $\mathcal{O}'$ in an obvious way. Since $\mathbf{T}_{\mathcal{D}}$ is reduced $\pi(\eta) \neq 0$.

One can also verify that

$$\pi(\eta) = \langle \eta, \eta \rangle \tag{2.43}$$

up to a unit in $\mathcal{O}$.

We will say that $\mathcal{D}_1 \supset \mathcal{D}$ if we obtain $\mathcal{D}_1$ by relaxing certain of the hypotheses on $\mathcal{D}$, i.e., if $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}, \mathcal{M})$ and $\mathcal{D}_1 = (\cdot, \Sigma_1, \mathcal{O}_1, \mathcal{M}_1)$ we allow that $\Sigma_1 \supset \Sigma$, any $\mathcal{O}_1$, $\mathcal{M} \supset \mathcal{M}_1$ (but of the same type) and if $\cdot$ is Se or str in $\mathcal{D}$ it can be Se, str, ord or unrestricted in $\mathcal{D}_1$, if $\cdot$ is fl in $\mathcal{D}_1$ it can be fl or unrestricted in $\mathcal{D}_1$. We use the term *restricted* to signify that $\cdot$ is Se, str, fl or ord. The following theorem reduces conjecture 2.16 to a 'class number' criterion. For an interpretation of the right-hand side of the inequality in the theorem as the order of a cohomology group, see Proposition 1.2. For an interpretation of the left-hand side in terms of the value of an inner product, see Proposition 4.4.

THEOREM 2.17. *Assume, as above, that $\rho_{f,\lambda}$ is a deformation of ρ_0 of type $\mathcal{D} = (\cdot, \Sigma, \mathcal{O} = \mathcal{O}_{f,\lambda}, \mathcal{M})$ with* $\cdot =$ Se, str *or* fl. *Suppose that*

$$\#\mathcal{O}/\pi(\eta_{\mathcal{D},f}) \geq \#\mathfrak{p}_{R,f}/\mathfrak{p}_{R,f}^2.$$

Then

(i) *$\varphi_{\mathcal{D}_1}: R_{\mathcal{D}_1} \simeq \mathbf{T}_{\mathcal{D}_1}$ is an isomorphism for all (restricted) $\mathcal{D}_1 \supset \mathcal{D}$.*

(ii) *$\mathbf{T}_{\mathcal{D}_1}$ is a complete intersection (over $\mathcal{O}_1$ if $\cdot$ is* Se, str *or* fl) *for all restricted $\mathcal{D}_1 \supset \mathcal{D}$.*

Proof. Let us write $\mathbf{T}$ for $\mathbf{T}_{\mathcal{D}}$, $\mathfrak{p}_{\mathbf{T}}$ for $\mathfrak{p}_{\mathbf{T},f}$, $\mathfrak{p}_R$ for $\mathfrak{p}_{R,f}$ and η for $\eta_{\mathcal{D},f}$. Then we always have

$$\#\mathcal{O}/\eta \leq \#\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2. \tag{2.44}$$

(Here and in what follows we sometimes write η for $\pi(\eta)$ if the context makes this reasonable.) This is proved as follows. $\mathbf{T}/\eta$ acts faithfully on $\mathfrak{p}_{\mathbf{T}}$. Hence the Fitting ideal of $\mathfrak{p}_{\mathbf{T}}$ as a $\mathbf{T}/\eta$-module is zero. The same is then true of $\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2$ as an $\mathcal{O}/\eta = (\mathbf{T}/\eta)/\mathfrak{p}_{\mathbf{T}}$-module. So the Fitting ideal of $\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2$ as an $\mathcal{O}$-module is contained in (η) and the conclusion is then easy. So together with the hypothesis of the theorem we get inequalities (and hence equalities)

$$\#\,\mathcal{O}/\pi(\eta) \geq \#\mathfrak{p}_R/\mathfrak{p}_R^2 \geq \#\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2 \geq \#\mathcal{O}/\pi(\eta).$$

By Proposition 2 of the appendix $\mathbf{T}$ is a complete intersection over $\mathcal{O}$. Part (ii) of the theorem then follows for $\mathcal{D}$. Part (i) follows for $\mathcal{D}$ from Proposition 1 of the appendix.

We now prove inductively that we can deduce the same inequality

$$\#\,\mathcal{O}_1/\eta_{\mathcal{D}_1,f} \geq \#\mathfrak{p}_{R_1,f}/\mathfrak{p}_{R_1,f}^2 \tag{2.45}$$

for $\mathcal{D}_1 \supset \mathcal{D}$ and $R_1 = R_{\mathcal{D}_1}$. The above argument will then prove the theorem for $\mathcal{D}_1$. We explain this first in the case $\mathcal{D}_1 = \mathcal{D}_q$ where $\mathcal{D}_q$ differs from $\mathcal{D}$ only in replacing Σ by $\Sigma \cup \{q\}$. Let us write $\mathbf{T}_q$ for $\mathbf{T}_{\mathcal{D}_q}$, $\mathfrak{p}_{R,q}$ for $\mathfrak{p}_{R,f}$ with $R = R_{\mathcal{D}_q}$ and η_q for $\eta_{\mathcal{D}_q,f}$. We recall that $U_q = 0$ in $\mathbf{T}_q$.

We choose isomorphisms

$$\mathbf{T} \simeq \mathrm{Hom}_{\mathcal{O}}(\mathbf{T}, \mathcal{O}), \qquad \mathbf{T}_q \simeq \mathrm{Hom}_{\mathcal{O}}(\mathbf{T}_q, \mathcal{O}) \tag{2.46}$$

coming from the fact that each of the rings is a Gorenstein $\mathcal{O}$-algebra. If $\alpha_q: \mathbf{T}_q \to \mathbf{T}$ is the natural map we may consider the element $\Delta_q = \alpha_q \circ \hat{\alpha}_q \in \mathbf{T}$ where the adjoint is with respect to the above isomorphisms. Then it is clear that

$$\Big(\alpha_q(\eta_q)\Big) = (\eta\Delta_q) \tag{2.47}$$

as principal ideals of $\mathbf{T}$. In particular $\pi(\eta_q) = \pi(\eta\,\Delta_q)$ in $\mathcal{O}$.

Now it follows from Proposition 2.7 that the principal ideal (Δ_q) is given by

$$(\Delta_q) = \Big((q-1)^2(T_q^2 - \langle q\rangle(1+q)^2)\Big). \tag{2.48}$$

In the statement of Proposition 2.7 we used $\mathbf{Z}_p$-pairings

$$\mathbf{T} \simeq \mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{T}, \mathbf{Z}_p), \qquad \mathbf{T}_q \simeq \mathrm{Hom}_{\mathbf{Z}_p}(\mathbf{T}_q, \mathbf{Z}_p)$$

to define $(\Delta_q) = (\alpha_q \circ \hat{\alpha}_q)$. However, using the description of the pairings as $W(k_{\mathfrak{m}})$-algebras derived from these $\mathbf{Z}_p$-pairings in the paragraph following (2.42) we see that the ideal (Δ_q) is unchanged when we use $W(k_{\mathfrak{m}})$-algebra pairings, and hence also when we extend scalars to $\mathcal{O}$ as in (2.42).

On the other hand

$$\#\mathfrak{p}_{R,q}/\mathfrak{p}_{R,q}^2 \le \#\mathfrak{p}_R/\mathfrak{p}_R^2 \cdot \#\left\{\mathcal{O}/(q-1)^2\left(T_q^2 - \langle q\rangle(1+q)^2\right)\right\}$$

by Propositions 1.2 and 1.7. Combining this with (2.47) and (2.48) gives (2.45).

If $\mathcal{M} \neq \phi$ we use a similar argument to pass from $\mathcal{D}$ to $\mathcal{D}_q$ where this time $\mathcal{D}_q$ signifies that $\mathcal{D}$ is unchanged except for dropping q from $\mathcal{M}$. In each of types (A), (B), and (C) one checks from Propositions 1.2 and 1.8 that

$$\#\mathfrak{p}_{R,q}/\mathfrak{p}_{R,q}^2 \le \#\mathfrak{p}_R/\mathfrak{p}_R^2 \cdot \#H^0(\mathbf{Q}_q, V^*).$$

This is in agreement with Propositions 2.10, 2.12 and 2.13 which give the corresponding change in η by the method described above.

To change from an $\mathcal{O}$-algebra to an $\mathcal{O}_1$-algebra is straightforward (the complete intersection property can be checked using [Ku1, Cor. 2.8 on p. 209]), and to change from Se to ord we use (1.4) and (2.32). The change from str to ord reduces to this since by Proposition 1.1 strict deformations and Selmer deformations are the same. Note that for the ord case if R is a local Noetherian ring and $f \in R$ is not a unit and not a zero divisor, then R is a complete intersection if and only if R/f is (cf. [BH, Th. 2.3.4]). This completes the proof of the theorem. □

Remark 2.18. If we suppose in the Selmer case that f has level N with $p \nmid N$ we can also consider the ring $\mathbf{T}_H(M_0)_{\mathfrak{m}_0}$ (with M_0 as in (2.24) and $\mathfrak{m}_0$ defined in the same way as for $\mathbf{T}_H(M)$). This time set

$$T_0 = \mathbf{T}_H(M_0)_{\mathfrak{m}_0} \underset{W(k_{\mathfrak{m}_0})}{\otimes} \mathcal{O}, \qquad T = \mathbf{T}_H(M)_{\mathfrak{m}} \underset{W(k_{\mathfrak{m}})}{\otimes} \mathcal{O}.$$

Define $\eta_0, \eta, \mathfrak{p}_0$ and $\mathfrak{p}$ with respect to these rings, and let $(\Delta_p) = \alpha_p \circ \hat{\alpha}_p$ where $\alpha_p : T \to T_0$ and the adjoint is taken with respect to $\mathcal{O}$-pairings on T and T_0. We then have by Proposition 2.4

$$(2.49) \qquad (\eta_p) = (\eta \cdot \Delta_p) = \left(\eta \cdot \left(T_p^2 - \langle p\rangle(1+p)^2\right)\right) = \left(\eta \cdot (a_p^2 - \langle p\rangle)\right)$$

as principal ideals of T, where a_p is the unit root of $x^2 - T_p x + p\langle p\rangle = 0$.

Remark. For some earlier work on how deformation rings change with Σ see [Bo].

Chapter 3

In this chapter we prove the main results about Conjecture 2.16. We begin by showing that the bound for the Selmer group to which it was reduced in Theorem 2.17 can be checked if one knows that the minimal Hecke ring is a complete intersection. Combining this with the main result of [TW] we complete the proof of Conjecture 2.16 under a hypothesis that ensures that a minimal Hecke ring exists.

Estimates for the Selmer group

Let ρ_0: $\mathrm{Gal}(\mathbf{Q}_\Sigma/\mathbf{Q}) \to \mathrm{GL}_2(k)$ be an odd irreducible representation which we will assume is modular. Let $\mathcal{D}$ be a deformation theory of type $(\cdot, \Sigma, \mathcal{O}, \mathcal{M})$ such that ρ_0 is type $\mathcal{D}$, where $\cdot$ is Selmer, strict or flat. We remind the reader that k is assumed to be the residue field of $\mathcal{O}$. Then as explained in Theorem 2.14, we can pick a modular lifting $\rho_{f,\lambda}$ of ρ_0 of type $\mathcal{D}$ (altering k if necessary and replacing $\mathcal{O}$ by a ring containing $\mathcal{O}_{f,\lambda}$) provided that ρ_0 is not induced from a character of $\mathbf{Q}(\sqrt{-3})$ if $p = 3$. For the rest of this chapter, we will make the assumption that ρ_0 is not of this exceptional type. Theorem 2.14 also specifies a certain minimum level and character for f and in particular ensures that we can pick f to have level prime to p when $\rho_0|_{D_p}$ is associated to a finite flat group scheme over $\mathbf{Z}_p$ and $\det \rho_0|_{I_p} = \omega$.

In Chapter 2, Section 3, we defined a ring $\mathbf{T}_\mathcal{D}$ associated to $\mathcal{D}$. Here we make a slight modification of this ring. In the case where $\cdot$ is Selmer and $\rho_0|_{D_p}$ is associated to a finite flat group scheme and $\det \rho_0|_{I_p} = \omega$ we set

$$\mathbf{T}_{\mathcal{D}_0} = \mathbf{T}'_H(M_0)_{\mathfrak{m}'_0} \underset{W(k_0)}{\otimes} \mathcal{O} \tag{3.1}$$

with M_0 as in (2.24), H defined following (2.24) (it is actually a subgroup of $(\mathbf{Z}/M_0\mathbf{Z})^*$) and $\mathfrak{m}'_0$ the maximal ideal of $\mathbf{T}'_H(M_0)$ associated to ρ_0. The same proof as in Proposition 2.15 ensures that there is a maximal ideal $\mathfrak{m}_0$ of $\mathbf{T}_H(M_0)$ with $\mathfrak{m}_0 \cap \mathbf{T}'_H(M_0) = \mathfrak{m}'_0$ and such that the natural map

$$\mathbf{T}_{\mathcal{D}_0} = \mathbf{T}'_H(M_0)_{\mathfrak{m}'_0} \underset{W(k_0)}{\otimes} \mathcal{O} \to \mathbf{T}_H(M_0)_{\mathfrak{m}_0} \underset{W(k_0)}{\otimes} \mathcal{O} \tag{3.2}$$

is an isomorphism. The maximal ideal $\mathfrak{m}_0$ which we choose is characterized by the properties that $\rho_{\mathfrak{m}_0} = \rho_0$ and $U_q \in \mathfrak{m}_0$ for $q \in \Sigma - \mathcal{M} \cup \{p\}$. (The value of T_p or of U_q for $q \in \mathcal{M}$ is determined by the other operators; see the proof of Proposition 2.15.) We now define $\mathbf{T}_{\mathcal{D}_0}$ in general by the following:

$$\mathbf{T}_{\mathcal{D}_0} \text{ is given by (3.1)} \quad \text{if } \cdot \text{ is Se and } \rho_0|_{D_p} \text{ is associated to a finite flat group scheme over } \mathbf{Z}_p \text{ and } \det \rho_0|_{I_p} = \omega;$$

(3.3)

$$\mathbf{T}_{\mathcal{D}_0} = \mathbf{T}_{\mathcal{D}} \quad \text{if } \cdot \text{ is str or fl, or } \rho_0|_{D_p} \text{ is not associated to a finite flat group scheme over } \mathbf{Z}_p, \text{ or } \det \rho_0|_{I_p} \neq \omega.$$

We choose a pair (f, λ) of minimum level and character as given by Theorem 2.14 and this gives a homomorphism of $\mathcal{O}$-algebras

$$\pi_f : \mathbf{T}_{\mathcal{D}_0} \to \mathcal{O} \supseteq \mathcal{O}_{f,\lambda}.$$

We set $\mathfrak{p}_{\mathbf{T},f} = \ker \pi_f$ and similarly we let $\mathfrak{p}_{R,f}$ denote the inverse image of $\mathfrak{p}_{\mathbf{T},f}$ in $R_{\mathcal{D}}$. We define a principal ideal $(\eta_{\mathbf{T},f})$ of $\mathbf{T}_{\mathcal{D}_0}$ by taking an adjoint $\hat{\pi}_f$ of π_f with respect to pairings as in (2.42) and write

$$\eta_{\mathbf{T},f} = (\hat{\pi}_f(1)).$$

Note that $\mathfrak{p}_{\mathbf{T},f}/\mathfrak{p}^2_{\mathbf{T},f}$ is finite and $\pi_f(\eta_{\mathbf{T},f}) \neq 0$ because $\mathbf{T}_{\mathcal{D}_0}$ is reduced. We also write $\eta_{\mathbf{T},f}$ for $\pi_f(\eta_{\mathbf{T},f})$ if the context makes this usage reasonable. We let $V_f = \operatorname{Ad} \rho_{\mathfrak{p}} \otimes_{\mathcal{O}} K/\mathcal{O}$ where $\rho_{\mathfrak{p}}$ is the extension of scalars of $\rho_{f,\lambda}$ to $\mathcal{O}$.

THEOREM 3.1. *Assume that $\mathcal{D}$ is minimal, i.e., $\sum = \mathcal{M} \cup \{p\}$, and that ρ_0 is absolutely irreducible when restricted to $\mathbf{Q}\left(\sqrt{(-1)^{\frac{p-1}{2}} p}\right)$. Then*

(i) $\# H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_f) \leq \#(\mathfrak{p}_{\mathbf{T},f}/\mathfrak{p}^2_{\mathbf{T},f})^2 \cdot c_p/\#(\mathcal{O}/\eta_{\mathbf{T},f})$

where $c_p = \#(\mathcal{O}/U_p^2 - \langle p \rangle) < \infty$ when ρ_0 is Selmer and $\rho_0|_{D_p}$ is associated to a finite flat group scheme over $\mathbf{Z}_p$ and $\det \rho_0|_{I_p} = \omega$, and $c_p = 1$ otherwise;

(ii) *if $\mathbf{T}_{\mathcal{D}_0}$ is a complete intersection over $\mathcal{O}$ then* (i) *is an equality, $R_{\mathcal{D}} \simeq \mathbf{T}_{\mathcal{D}}$ and $\mathbf{T}_{\mathcal{D}}$ is a complete intersection.*

In general, for any (not necessarily minimal) $\mathcal{D}$ of Selmer, strict or flat type, and any $\rho_{f,\lambda}$ of type $\mathcal{D}$, $\# H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_f) < \infty$ if ρ_0 is as above.

Remarks. The finiteness was proved by Flach in [Fl] under some restrictions on f, p and $\mathcal{D}$ by a different method. In particular, he did not consider the strict case. The bound we obtain in (i) is in fact the actual order of $H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_f)$ as follows from the main result of [TW] which proves the hypothesis of part (ii). Then applying Theorem 2.17 we obtain the order of this group for more general $\mathcal{D}$'s associated to ρ_0 under the condition that a minimal $\mathcal{D}$ exists associated to ρ_0. This is stated in Theorem 3.3.

The case where the projective representation associated to ρ_0 is dihedral does not always have the property that a twist of it has an associated minimal $\mathcal{D}$. In the case where the associated quadratic field is imaginary we will give a different argument in Chapter 4.

Proof. We will assume throughout the proof that $\mathcal{D}$ is minimal, indicating only at the end the slight changes needed for the final assertion of the theorem. Let Q be a finite set of primes disjoint from Σ satisfying $q \equiv 1(p)$ and $\rho_0(\mathrm{Frob}\, q)$ having distinct eigenvalues for each $q \in Q$. For the minimal deformation problem $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}, \mathcal{M})$, let $\mathcal{D}_Q$ be the deformation problem described before (2.34); i.e., it is the refinement of $(\cdot, \Sigma \cup Q, \mathcal{O}, \mathcal{M})$ obtained by imposing the additional restriction (2.34) at each $q \in Q$. (We will assume for the proof that $\mathcal{O}$ is chosen so $\mathcal{O}/\lambda = k$ contains the eigenvalues of $\rho_0(\mathrm{Frob}\, q)$ for each $q \in Q$.) We set

$$\mathbf{T} = \mathbf{T}_{\mathcal{D}_0}, \qquad R = R_{\mathcal{D}}$$

and recall the definition of $\mathbf{T}_Q$ and R_Q from Chapter 2, §3 (cf. (2.35)). We write V for V_f and recall the definition of $V_{(q)}$ following (2.38). Also remember that $\mathfrak{m}_Q$ is a maximal ideal of $\mathbf{T}_H(Nq_1 \dots q_r)$ as in (2.35) for which $\rho_{\mathfrak{m}_Q} \simeq \rho_0$ over $\bar{\mathbf{F}}_p$ (recall that this uses the same choice of embedding $k_{\mathfrak{m}_Q} \longrightarrow k$ as in the definition of $\mathbf{T}_Q$). We use $\mathfrak{m}_Q$ also to denote the maximal ideal of $\mathbf{T}_Q$ if the context makes this reasonable.

Consider the exact and commutative diagram

$$\begin{array}{ccccccc}
0 \longrightarrow & H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V) & \longrightarrow & H^1_{\mathcal{D}_Q}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V) & \xrightarrow{\delta_Q} & \prod_{q\in Q} H^1(\mathbf{Q}_q^{\mathrm{unr}}, V^{(q)})^{\mathrm{Gal}(\mathbf{Q}_q^{\mathrm{unr}}/\mathbf{Q}_q)} & \\
& \wr\| & & \wr\| & & & \\
0 \longrightarrow & (\mathfrak{p}_R/\mathfrak{p}_R^2)^* & \longrightarrow & (\mathfrak{p}_{R_Q}/\mathfrak{p}_{R_Q}^2)^* & & \Big\uparrow \iota_Q & \\
& \uparrow & & \uparrow & & & \\
0 \longrightarrow & (\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2)^* & \longrightarrow & (\mathfrak{p}_{\mathbf{T}_Q}/\mathfrak{p}_{\mathbf{T}_Q}^2)^* & \xrightarrow{u_Q} & K_Q \longrightarrow 0 &
\end{array}$$

where K_Q is by definition the cokernel in the horizontal sequence and $*$ denotes $\mathrm{Hom}_{\mathcal{O}}(\ , K/\mathcal{O})$ for K the field of fractions of $\mathcal{O}$. The key result is:

LEMMA 3.2. *The map ι_Q is injective for any finite set of primes Q satisfying*

$$q \equiv 1(p), T_q^2 \not\equiv \langle q \rangle (1+q)^2 \bmod \mathfrak{m} \textit{ for all } q \in Q.$$

Proof. Note that the hypotheses of the lemma ensure that $\rho_0(\mathrm{Frob}\, q)$ has distinct eigenvalues for each $q \in Q$. First, consider the ideal $\mathfrak{a}_Q$ of R_Q defined

by

$$(3.4)\qquad \mathfrak{a}_Q=\left\{a_i-1, b_i, c_i, d_i-1 : \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}=\rho_{\mathcal{D}_Q}(\sigma_i) \text{ with } \sigma_i \in I_{q_i}, q_i \in Q\right\}.$$

Then the universal property of R_Q shows that $R_Q/\mathfrak{a}_Q \simeq R$. This permits us to identify $(\mathfrak{p}_R/\mathfrak{p}_R^2)^*$ as

$$(\mathfrak{p}_R/\mathfrak{p}_R^2)^* = \{f \in (\mathfrak{p}_{R_Q}/\mathfrak{p}_{R_Q}^2)^* : f(\mathfrak{a}_Q) = 0\}.$$

If we prove the same relation for the Hecke rings, i.e., with $\mathbf{T}$ and $\mathbf{T}_Q$ replacing R and R_Q then we will have the injectivity of ι_Q. We will write $\bar{\mathfrak{a}}_Q$ for the image of $\mathfrak{a}_Q$ in $\mathbf{T}_Q$ under the map φ_Q of (2.37).

It will be enough to check that for any $q \in Q'$, Q' a subset of Q, $\mathbf{T}_{Q'}/\bar{\mathfrak{a}}_q \simeq \mathbf{T}_{Q'-\{q\}}$ where $\mathfrak{a}_q$ is defined as in (3.4) but with Q replaced by q. Let $N' = N(\rho_0)p^{\delta(\rho_0)} \cdot \prod_{q_i \in Q'-\{q\}} q_i$ where $\delta(\rho_0)$ is as defined in Theorem 2.14. Then take an element $\sigma \in I_q \subseteq \mathrm{Gal}(\bar{\mathbf{Q}}_q/\mathbf{Q}_q)$ which restricts to a generator of $\mathrm{Gal}(\mathbf{Q}(\zeta_{N'q})/\mathbf{Q}(\zeta_{N'}))$. Then $\det(\sigma) = \langle t_q \rangle \in \mathbf{T}_{Q'}$ in the representation to $\mathrm{GL}_2(\mathbf{T}_{Q'})$ defined after Theorem 2.1. (Thus $t_q \equiv 1(N')$ and t_q is a primitive root mod q.) It is easily checked that

$$(3.5)\qquad J_H(N', q)_{\mathfrak{m}_{Q'}}(\bar{\mathbf{Q}}) \simeq J_H(N'q)_{\mathfrak{m}_{Q'}}(\bar{\mathbf{Q}})\,[\langle t_q \rangle - 1].$$

Here H is still a subgroup of $(\mathbf{Z}/M_0\mathbf{Z})^*$. (We use here that ρ_0 is not reducible for the injectivity and also that ρ_0 is not induced from a character of $\mathbf{Q}(\sqrt{-3})$ for the surjectivity when $p = 3$. The latter is to avoid the ramification points of the covering $X_H(N'q) \to X_H(N', q)$ of order 3 which can give rise to invariant divisors of $X_H(N'q)$ which are not the images of divisors on $X_H(N', q)$.)

Now by Corollary 1 to Theorem 2.1 the Pontrjagin duals of the modules in (3.5) are free of rank two. It follows that

$$(3.6)\qquad (\mathbf{T}_H(N'q)_{\mathfrak{m}_{Q'}})^2/(\langle t_q \rangle - 1) \simeq (\mathbf{T}_H(N', q)_{\mathfrak{m}_{Q'}})^2.$$

The hypotheses of the lemma imply the condition that $\rho_0(\mathrm{Frob}\, q)$ has distinct eigenvalues. So applying Proposition 2.4′ (at the end of §2) and the remark following it (or using the fact remarked in Chapter 2, §3 that this condition implies that ρ_0 does not occur as the residual representation associated to any form which has the special representation at q) we see that after tensoring over $W(k_{\mathfrak{m}_{Q'}})$ with $\mathcal{O}$, the right-hand side of (3.6) can be replaced by $\mathbf{T}^2_{Q'-\{q\}}$ thus giving

$$\mathbf{T}^2_{Q'}/\bar{\mathfrak{a}}_q \simeq \mathbf{T}^2_{Q'-\{q\}},$$

since $\langle t_q \rangle - 1 \in \bar{\mathfrak{a}}_q$. Repeated inductively this gives the desired relation $\mathbf{T}_Q/\bar{\mathfrak{a}}_Q \simeq \mathbf{T}$, and completes the proof of the lemma. □

Suppose now that Q is a finite set of primes chosen as in the lemma. Recall that from the theory of congruences (Prop. 2.4′ at the end of §2)

$$\eta_{\mathbf{T}_{Q,f}}/\eta_{\mathbf{T},f} = \prod_{q\in Q}(q-1),$$

the factors $(\alpha_q^2 - \langle q\rangle)$ being units by our hypotheses on $q \in Q$. (We only need that the right-hand side divides the left which is somewhat easier.) Also, from the theory of Fitting ideals (see the proof of (2.44))

$$\begin{aligned}\#(\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2) &\geq \#(\mathcal{O}/\eta_{\mathbf{T}_f}) \\ \#(\mathfrak{p}_{\mathbf{T}_Q}/\mathfrak{p}_{\mathbf{T}_Q}^2) &\geq \#(\mathcal{O}/\eta_{\mathbf{T}_{Q,f}}).\end{aligned}$$

We deduce that

$$\#K_Q \geq \#\left(\mathcal{O}\Big/\prod_{q\in Q}(q-1)\right) \cdot t^{-1}$$

where $t = \#(\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2)/\#(\mathcal{O}/\eta_{\mathbf{T},f})$. Since the range of ι_Q has order given by

$$\#\left\{\mathcal{O}\Big/\prod_{q\in Q}(q-1)\right\},$$

we compute that the index of the image of ι_Q is $\leq t$ as ι_Q is injective.

Keeping our assumption on Q from Lemma 3.2, consider the kernel of λ^M applied to the diagram at the beginning of the proof of the theorem. Then with M chosen large enough so that λ^M annihilates $\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2$ (which is finite because $\mathbf{T}$ is reduced) we get:

$$\begin{array}{ccccccc}
0 \longrightarrow H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V[\lambda^M]) & \longrightarrow & H^1_{\mathcal{D}_Q}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V[\lambda^M]) & \xrightarrow{\delta_Q} & \prod_{q\in Q} H^1(\mathbf{Q}_q^{\mathrm{unr}}, V^{(q)}[\lambda^M])^{\mathrm{Gal}(\mathbf{Q}_q^{\mathrm{unr}}/\mathbf{Q}_q)} & & \\
\uparrow & & \uparrow \psi_Q & & \uparrow \iota_Q & & \\
0 \longrightarrow (\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2)^* & \longrightarrow & (\mathfrak{p}_{\mathbf{T}_Q}/\mathfrak{p}_{\mathbf{T}_Q}^2)^*\,[\lambda^M] & \longrightarrow & K_Q[\lambda^M] \to (\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2)^*. & &
\end{array}$$

See (1.7) for the justification that λ^M can be taken inside the parentheses in the first two terms. Let $X_Q = \psi_Q((\mathfrak{p}_{\mathbf{T}_Q}/\mathfrak{p}_{\mathbf{T}_Q}^2)^*\,[\lambda^M])$. Then we can estimate the order of $\delta_Q(X_Q)$ using the fact that the image of ι_Q has index at most t. We get

$$\#\delta_Q(X_Q) \geq \left(\prod_{q\in Q}\#\mathcal{O}/(\lambda^M, q-1)\right)\cdot(1/t)\cdot(1/\#(\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}_{\mathbf{T}}^2)). \tag{3.7}$$

Now we choose Q to be a set of primes with the property that

$$\varepsilon_Q : H^1_{\mathcal{D}^*}(\mathbf{Q}_\Sigma/\mathbf{Q}, V^*_{\lambda^M}) \to \prod_{q\in Q} H^1(\mathbf{Q}_q, V^*_{\lambda^M}) \tag{3.8}$$

is injective. We also keep the condition that ι_Q is injective by only allowing Q to contain primes of the form given in the lemma. In addition, we require these q's to satisfy $q \equiv 1(p^M)$.

To see that this can be done, suppose that $x \in \ker \varepsilon_Q$ and $\lambda x = 0$ but $x \neq 0$. We have a commutative diagram

$$\begin{array}{ccc} H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, V^*_{\lambda^M})\,[\lambda] & \xrightarrow{\varepsilon_Q} & \prod_{q\in Q} H^1(\mathbf{Q}_q, V^*_{\lambda^M})\,[\lambda] \\ \wr\| & & \wr\| \\ H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, V^*_\lambda) & \xrightarrow{\bar\varepsilon_Q} & \prod_{q\in Q} H^1(\mathbf{Q}_q, V^*_\lambda) \end{array}$$

the right-hand isomorphisms coming from our particular choices of q's and the left-hand isomorphism from our hypothesis on ρ_0. The same diagram will hold if we replace Q by $Q_0 = Q \cup \{q_0\}$ and we now need to show that we can choose q_0 so that $\bar\varepsilon_{Q_0}(x) \neq 0$.

The restriction map

$$H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, V^*_\lambda) \to \mathrm{Hom}(\mathrm{Gal}(\bar{\mathbf{Q}}/K_0(\zeta_p)), V^*_\lambda)^{\mathrm{Gal}(K_0(\zeta_p)/\mathbf{Q})}$$

has kernel $H^1(K_0(\zeta_p)/\mathbf{Q}, k(1))$ by Proposition 1.11 where here K_0 is the splitting field of ρ_0. Now if $x \in H^1(K_0(\zeta_p)/\mathbf{Q}, k(1))$ and $x \neq 0$ then $p = 3$ and x factors through an abelian extension L of $\mathbf{Q}(\zeta_3)$ of exponent 3 which is non-abelian over $\mathbf{Q}$. In this exceptional case, L must ramify at some prime $\mathfrak{q}$ of $\mathbf{Q}(\zeta_3)$, and if $\mathfrak{q}$ lies over the rational prime $q \neq 3$ then the composite map

$$H^1(K_0(\zeta_3)/\mathbf{Q}, k(1)) \longrightarrow H^1(\mathbf{Q}_q^{\mathrm{unr}}, k(1)) \longrightarrow H^1(\mathbf{Q}_q^{\mathrm{unr}}, (\mathcal{O}/\lambda^M)(1))$$

is nonzero on x. But then x is not of type $\mathcal{D}^*$ which gives a contradiction. This only leaves the possibility that $L = \mathbf{Q}(\zeta_3, \sqrt[3]{3})$ but again this means that x is not of type $\mathcal{D}^*$ as locally at the prime above 3, L is not generated by the cube root of a unit over $\mathbf{Q}_3(\zeta_3)$. This argument holds whether or not $\mathcal{D}$ is minimal.

So x, which we view in $\ker \bar\varepsilon_Q$, gives a nontrivial Galois-equivariant homomorphism $f_x \in \mathrm{Hom}(\mathrm{Gal}(\bar{\mathbf{Q}}/K_0(\zeta_p)), V^*_\lambda)$ which factors through an abelian extension M_x of $K_0(\zeta_p)$ of exponent p. Specifically we choose M_x to be the minimal such extension. Assume first that the projective representation $\tilde\rho_0$ associated to ρ_0 is not dihedral so that $\mathrm{Sym}^2 \rho_0$ is absolutely irreducible. Pick a $\sigma \in \mathrm{Gal}(M_x(\zeta_{p^M})/\mathbf{Q})$ satisfying

(3.9)
(i) $\rho_0(\sigma)$ has order $m \geq 3$ with $(m, p) = 1$,
(ii) σ fixes $\mathbf{Q}(\det \rho_0)\,(\zeta_{p^M})$,
(iii) $f_x(\sigma^m) \neq 0$.

To show that this is possible, observe first that the first two conditions can be achieved by Lemma 1.10(i) and the subsequent remark. Let σ_1 be an el-

ement satisfying (i) and (ii) and let $\bar{\sigma}_1$ denote its image in $\mathrm{Gal}(K_0(\zeta_p)/\mathbf{Q})$. Then $\langle\bar{\sigma}_1\rangle$ acts on $G = \mathrm{Gal}(M_x/K_0(\zeta_p))$ and under this action G decomposes as $G \simeq G_1 \oplus G_1'$ where σ_1 acts trivially on G_1 and without fixed points on G_1'. If X is any irreducible Galois stable $\bar{k}$-subspace of $f_x(G) \otimes_{\mathbf{F}_p} \bar{k}$ then $\ker(\sigma_1 - 1)|_X \neq 0$ since $\mathrm{Sym}^2 \rho_0$ is assumed absolutely irreducible. So also $\ker(\sigma_1 - 1)|_{f_x(G)} \neq 0$ and thus we can find $\tau \in G_1$ such that $f_x(\tau) \neq 0$. Viewing τ as an element of G we then take

$$\tau_1 = \tau \times 1 \in \mathrm{Gal}(M_x(\zeta_{p^M})/K_0(\zeta_p)) \simeq G \times \mathrm{Gal}(K_0(\zeta_{p^M})/K_0(\zeta_p))$$

(This decomposition holds because M_x is minimal and because $\mathrm{Sym}^2 \rho_0$ and μ_p are distinct from the trivial representation.) Now τ_1 commutes with σ_1 and either $f_x((\tau_1 \sigma_1)^m) \neq 0$ or $f_x(\sigma_1^m) \neq 0$. Since $\rho_0(\tau_1\sigma_1) = \rho_0(\sigma_1)$ this gives (3.9) with at least one of $\sigma = \tau_1\sigma_1$ or $\sigma = \sigma_1$. We may then choose q_0 so that $\mathrm{Frob}\, q_0 = \sigma$ and we will then have $\bar{\varepsilon}_{Q_0}(x) \neq 0$. Note that conditions (i) and (ii) imply that $q_0 \equiv 1(p)$ and also that $\rho_0(\sigma)$ has distinct eigenvalues, thus giving both the hypotheses of Lemma 3.2.

If on the other hand $\tilde{\rho}_0$ is dihedral then we pick σ's satisfying

(i) $\tilde{\rho}_0(\sigma) \neq 1$,

(ii) σ fixes $\mathbf{Q}(\zeta_{p^M})$,

(iii) $f_x(\sigma^m) \neq 0$,

with m the order of $\rho_0(\sigma)$ (and $p \nmid m$ since $\tilde{\rho}_0$ is dihedral). The first two conditions can be achieved using Lemma 1.12 and, in addition, we can assume that σ takes the eigenvalue 1 on any given irreducible Galois stable subspace X of $W_\lambda \otimes \bar{k}$. Arguing as above, we find a $\tau \in G_1$ such that $f_x(\tau) \neq 0$ and we proceed as before. Again, conditions (i) and (ii) imply the hypotheses of Lemma 3.2. So by successively adjoining q's we can assume that Q is chosen so that ε_Q is injective.

We have thus shown that we can choose $Q = \{q_1, \dots, q_r\}$ to be a finite set of primes $q_i \equiv 1(p^M)$ satisfying the hypotheses of Lemma 3.2 as well as the injectivity of ε_Q in (3.8). By Proposition 1.6, the injectivity of ε_Q implies that

(3.10) $$\# H^1_{\mathcal{D}}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V_f[\lambda^M]) = h_\infty \cdot \prod_{q \in \Sigma \cup Q} h_q.$$

Here we are using the convention explained after Proposition 1.6 to define $H^1_{\mathcal{D}}$. Now, as $\mathcal{D}$ was chosen to be minimal, $h_q = 1$ for $q \in \sum - \{p\}$ by Proposition 1.8. Also, $h_q = \#(\mathcal{O}/\lambda^M)^2$ for $q \in Q$. If $\cdot$ is str or fl then $h_\infty h_p = 1$ by Proposition 1.9 (iv) and (v). If $\cdot$ is Se, $h_\infty h_p \leq c_p$ by Proposition 1.9 (iii). (To compute this we can assume that I_p acts on W^0_λ via ω, as otherwise we

get $h_\infty h_p \le 1$. Then with this hypothesis, $(W^0_{\lambda^n})^*$ is easily verified to be unramified with $\mathrm{Frob}\, p$ acting as $U_p^2\langle p\rangle^{-1}$ by the description of $\rho_{f,\lambda}|_{D_p}$ in [Wil, Th. 2.1.4].) On the other hand, we have constructed classes which are ramified at primes in Q in (3.7). These are of type $\mathcal{D}_Q$. We also have classes in

$$\mathrm{Hom}(\mathrm{Gal}(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}), \mathcal{O}/\lambda^M) = H^1(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, \mathcal{O}/\lambda^M) \hookrightarrow H^1(\mathbf{Q}_{\Sigma\cup Q}/\mathbf{Q}, V_{\lambda^M})$$

coming from the cyclotomic extension $\mathbf{Q}(\zeta_{q_1} \dots \zeta_{q_r})$. These are of type $\mathcal{D}$ and disjoint from the classes obtained from (3.7). Combining these with (3.10) gives

$$\# H^1_{\mathcal{D}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_f[\lambda^M]) \le t \cdot \# (\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}^2_{\mathbf{T}}) \cdot c_p$$

as required. This proves part (i) of Theorem 3.1.

Now if we assume that $\mathbf{T}$ is a complete intersection we have that $t = 1$ by Proposition 2 of the appendix. In the strict or flat cases (and indeed in all cases where $c_p = 1$) this implies that $R_{\mathcal{D}} \simeq \mathbf{T}_{\mathcal{D}}$ by Proposition 1 of the appendix together with Proposition 1.2. In the Selmer case we get

$$(3.11) \qquad \# (\mathfrak{p}_{\mathbf{T}}/\mathfrak{p}^2_{\mathbf{T}}) \cdot c_p = \# (\mathcal{O}/\eta_{\mathbf{T},f})c_p = \# (\mathcal{O}/\eta_{\mathbf{T}_{\mathcal{D}},f}) \le \# (\mathfrak{p}_{\mathbf{T}_{\mathcal{D}}}/\mathfrak{p}^2_{\mathbf{T}_{\mathcal{D}}})$$

where the central equality is by Remark 2.18 and the right-hand inequality is from the theory of Fitting ideals. Now applying part (i) we see that the inequality in (3.11) is an equality. By Proposition 2 of the appendix, $\mathbf{T}_{\mathcal{D}}$ is also a complete intersection.

The final assertion of the theorem is proved in exactly the same way on noting that we only used the minimality to ensure that the h_q's were 1. In general, they are bounded independent of M and easily computed. (The only point to note is that if $\rho_{f,\lambda}$ is of multiplicative type at q then $\rho_{f,\lambda}|_{D_q}$ does not split.) □

Remark. The ring $\mathbf{T}_{\mathcal{D}_0}$ defined in (3.1) and used in this chapter should be the deformation ring associated to the following deformation problem $\mathcal{D}_0$. One alters $\mathcal{D}$ only by replacing the Selmer condition by the condition that the deformations be flat in the sense of Chapter 1, i.e., that each deformation ρ of ρ_0 to $\mathrm{GL}_2(A)$ has the property that for any quotient $A/\mathfrak{a}$ of finite order, $\rho|_{D_p} \bmod \mathfrak{a}$ is the Galois representation associated to the $\bar{\mathbf{Q}}_p$-points of a finite flat group scheme over $\mathbf{Z}_p$. (Of course, ρ_0 is ordinary here in contrast to our usual assumption for flat deformations.)

From Theorem 3.1 we deduce our main results about representations by using the main result of [TW], which proves the hypothesis of Theorem 3.1 (ii), and then applying Theorem 2.17. More precisely, the main result of [TW] shows that $\mathbf{T}$ is a complete intersection and hence that $t = 1$ as explained above. The hypothesis of Theorem 2.17 is then given by Theorem 3.1(i), together with the equality $t = 1$ (and the central equality of (3.11) in the

Selmer case) and Proposition 1.2. Strictly speaking, Theorem 1 of [TW] refers to a slightly smaller class of $\mathcal{D}$'s than those covered by Theorem 3.1 but up to a twist every such $\mathcal{D}$ is covered. It is straightforward to see that it is enough to check Theorem 3.3 for ρ_0 up to a suitable twist.

THEOREM 3.3. *Assume that ρ_0 is modular and absolutely irreducible when restricted to $\mathbf{Q}\left(\sqrt{(-1)^{\frac{p-1}{2}}p}\right)$. Assume also that ρ_0 is of type* (A), (B) *or* (C) *at each $q \neq p$ in Σ. Then the map $\varphi_\mathcal{D}$: $R_\mathcal{D} \longrightarrow \mathbf{T}_\mathcal{D}$ of Conjecture* 2.16 *is an isomorphism for all $\mathcal{D}$ associated to ρ_0, i.e., where $\mathcal{D} = (\cdot,\Sigma,\mathcal{O},\mathcal{M})$ with* $\cdot = \mathrm{Se}, \mathrm{str}, \mathrm{fl}$ *or* ord. *In particular if* $\cdot =$ Se, str *or* fl *and f is any newform for which $\rho_{f,\lambda}$ is a deformation of ρ_0 of type $\mathcal{D}$ then*

$$\# H^1_\mathcal{D}(\mathbf{Q}_\Sigma/\mathbf{Q}, V_f) = \#(\mathcal{O}/\eta_{\mathcal{D},f}) < \infty$$

where $\eta_{\mathcal{D},f}$ is the invariant defined in Chapter 2 *prior to* (2.43).

The condition at $q \neq p$ in Σ ensures that there is a minimal $\mathcal{D}$ associated to ρ_0. The computation of the Selmer group follows from Theorem 2.17 and Proposition 1.2. Theorem 0.2 of the introduction follows from Theorem 3.3, after it is checked that a twist of a ρ_0 as in Theorem 0.2 satisfies the hypotheses of Theorem 3.3.

Chapter 4

In this chapter we give a different (and slightly more general) derivation of the bound for the Selmer group in the CM case. In the first section we estimate the Selmer group using the main theorem of [Ru 4] which is based on Kolyvagin's method. In the second section we use a calculation of Hida to relate the η-invariant to special values of an L-function. Some of these computations are valid in the non-CM case also. They are needed if one wishes to give the order of the Selmer group in terms of the special value of an L-function.

1. The ordinary CM case

In this section we estimate the order of the Selmer group in the ordinary CM case. In Section 1 we use the proof of the main conjecture by Rubin to bound the Selmer group in terms of an L-function. The methods are standard (cf. [de Sh]) and some special cases have been described elsewhere (cf. [Guo]). In Section 2 we use a calculation of Hida to relate this to the η-invariant.

We assume that

$$\rho = \mathrm{Ind}_L^{\mathbf{Q}} \kappa : \ \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \ \rightarrow \ \mathrm{GL}_2(\mathcal{O}) \tag{4.1}$$

is the p-adic representation associated to a character $\kappa : \mathrm{Gal}(\bar{L}/L) \to \mathcal{O}^\times$ of an imaginary quadratic field L. We assume that p is unramified in L and that κ factors through an extension of L whose Galois group has the form $A \simeq \mathbf{Z}_p \oplus T$ where T is a finite group of order prime to p. The ring $\mathcal{O}$ is assumed to be the ring of integers of a local field with maximal ideal λ and we also assume that ρ is a Selmer deformation of $\rho_0 = \rho \bmod \lambda$ which is supposed irreducible with $\det \rho_0|_{I_p} = \omega$. In particular it follows that p splits in L, $p = \mathfrak{p}\bar{\mathfrak{p}}$ say, and that precisely one of κ, κ^* is ramified at $\mathfrak{p}$ (κ^* being the character $\tau \to \kappa(\sigma\tau\sigma^{-1})$ for any σ representing the nontrivial coset in $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \big/ \mathrm{Gal}(\bar{\mathbf{Q}}/L)$). We can suppose without loss of generality that κ is ramified at $\mathfrak{p}$.

We consider the representation module $V \simeq (K/\mathcal{O})^4$ (where K is the field of fractions of $\mathcal{O}$) and the representation is via $\mathrm{Ad}\,\rho$. In this case V splits as

$$V \simeq Y \oplus (K/\mathcal{O})(\psi) \oplus K/\mathcal{O}$$

where ψ is the quadratic character of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ associated to L. We let Σ denote a finite set of primes including all those which ramify in ρ (and in particular p). Our aim is to compute $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$. The decomposition of V gives a corresponding decomposition of $H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$ and we can use it to define $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y)$. Since $W^0 \subset Y$ (see Chapter 1 for the definition of W^0) we can define $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y)$ by

$$H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y) = \ker\{H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, Y) \to H^1(\mathbf{Q}_p^{\mathrm{unr}}, Y/W^0)\}.$$

Let Y^* be the arithmetic dual of Y, i.e., $\mathrm{Hom}(Y, \boldsymbol{\mu}_{p^\infty}) \otimes \mathbf{Q}_p/\mathbf{Z}_p$. Write ν for $\kappa\varepsilon/\kappa^*$ and let $L(\nu)$ be the splitting field of ν. Then we claim that $\mathrm{Gal}(L(\nu)/L) \simeq \mathbf{Z}_p \oplus T'$ with T' a finite group of order prime to p. For this it is enough to show that $\chi = \kappa\kappa^*/\varepsilon$ factors through a group of order prime to p since $\nu = \kappa^2\chi^{-1}$. Suppose that χ has order $m = m_0 p^r$ with $(m_0, p) = 1$. Then χ^{m_0} extends to a character of $\mathbf{Q}$ which is then unramified at p since the same is true of χ. Also it factors through an abelian extension of L with Galois group isomorphic to $\mathbf{Z}_p^2$ since χ factors through such an extension with Galois group isomorphic to $\mathbf{Z}_p^2 \oplus T_1$ with T_1 of order prime to p (the composite of the splitting fields of κ and κ^*). It follows that χ^{m_0} is also unramified outside p, whence it is trivial. This proves the claim.

Over L there is an isomorphism of Galois modules

$$Y^* \simeq (K/\mathcal{O})(\nu) \oplus (K/\mathcal{O})(\nu^{-1}\varepsilon^2).$$

In analogy to the above we define $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y^*)$ by

$$H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y^*) = \ker\{H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, Y^*) \to H^1(\mathbf{Q}_p^{\mathrm{unr}}, (W^0)^*)\}.$$

Analogous definitions apply if Y^* is replaced by $Y^*_{\lambda^n}$. Also we say informally that a cohomology class is Selmer at p if it vanishes in $H^1(\mathbf{Q}_p^{\mathrm{unr}}, (W^0)^*)$ (resp.

$H^1(\mathbf{Q}_p^{\mathrm{unr}}, (W^0_{\lambda^n})^*))$. Let M_∞ be the maximal abelian p-extension of $L(\nu)$ unramified outside $\mathfrak{p}$. The following proposition generalizes [CS, Prop. 5.9].

PROPOSITION 4.1. *There is an isomorphism*

$$H^1_{\mathrm{unr}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y^*) \xrightarrow{\sim} \mathrm{Hom}\,(\mathrm{Gal}(M_\infty/L(\nu)),\ (K/\mathcal{O})(\nu))^{\mathrm{Gal}(L(\nu)/L)}$$

where H^1_{unr} denotes the subgroup of classes which are Selmer at p and unramified everywhere else.

Proof. The sequence is obtained from the inflation-restriction sequence as follows. First we can replace $H^1(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y^*)$ by

$$\left\{H^1\left(\mathbf{Q}_\Sigma/L,\ (K/\mathcal{O})(\nu)\right) \oplus H^1\left(\mathbf{Q}_\Sigma/L,\ (K/\mathcal{O})(\nu^{-1}\varepsilon^2)\right)\right\}^\Delta$$

where $\Delta = \mathrm{Gal}(L/\mathbf{Q})$. The unramified condition then translates into the requirement that the cohomology class should lie in

$$\left\{H^1_{\mathrm{unr\ in\ }\Sigma-\mathfrak{p}}(\mathbf{Q}_\Sigma/L,\ (K/\mathcal{O})(\nu)) \oplus H^1_{\mathrm{unr\ in\ }\Sigma-\mathfrak{p}^*}\left(\mathbf{Q}_\Sigma/L,\ (K/\mathcal{O})(\nu^{-1}\varepsilon^2)\right)\right\}^\Delta.$$

Since Δ interchanges the two groups inside the parentheses it is enough to compute the first of them, i.e.,

$$H^1_{\mathrm{unr\ in\ }\Sigma-\mathfrak{p}}(\mathbf{Q}_\Sigma/L,\ K/\mathcal{O}(\nu)). \tag{4.2}$$

The inflation-restriction sequence applied to this gives an exact sequence

$$\begin{aligned}(4.3)\qquad 0 \quad \to \quad & H^1_{\mathrm{unr\ in\ }\Sigma-\mathfrak{p}}(L(\nu)/L,\ (K/\mathcal{O})(\nu)) \\ \to \quad & H^1_{\mathrm{unr\ in\ }\Sigma-\mathfrak{p}}(\mathbf{Q}_\Sigma/L,\ (K/\mathcal{O})(\nu)) \\ \to \quad & \mathrm{Hom}\,(\mathrm{Gal}(M_\infty/L(\nu)),\ (K/\mathcal{O})(\nu))^{\mathrm{Gal}(L(\nu)/L)}.\end{aligned}$$

The first term is zero as one easily checks using the divisibility of $(K/\mathcal{O})(\nu)$. Next note that $H^2(L(\nu)/L,\ (K/\mathcal{O})(\nu))$ is trivial. If $\nu \not\equiv 1(\lambda)$ this is straightforward (cf. Lemma 2.2 of [Ru1]). If $\nu \equiv 1(\lambda)$ then $\mathrm{Gal}\,(L(\nu)/L) \simeq \mathbf{Z}_p$ and so it is trivial in this case also. It follows that any class in the final term of (4.3) lifts to a class c in $H^1(\mathbf{Q}_\Sigma/L,\ (K/\mathcal{O})(\nu))$. Let L_0 be the splitting field of Y^*_λ. Then $M_\infty L_0/L_0$ is unramified outside $\mathfrak{p}$ and L_0/L has degree prime to p. It follows that c is unramified outside $\mathfrak{p}$. □

Now write $H^1_{\mathrm{str}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y^*_n)$ (where $Y^*_n = Y^*_{\lambda^n}$ and similarly for Y_n) for the subgroup of $H^1_{\mathrm{unr}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y^*_n)$ given by

$$H^1_{\mathrm{str}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y^*_n) = \left\{\alpha \in H^1_{\mathrm{unr}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y^*_n) :\ \alpha_p = 0 \text{ in } H^1(\mathbf{Q}_p,\ Y^*_n/(Y^*_n)^0)\right\}$$

where $(Y^*_n)^0$ is the first step in the filtration under D_p, thus equal to $(Y_n/Y^0_n)^*$ or equivalently to $(Y^*)^0_{\lambda^n}$ where $(Y^*)^0$ is the divisible submodule of Y^* on which the action of I_p is via ε^2. (If $p \neq 3$ one can characterize $(Y^*_n)^0$ as the

maximal submodule on which I_p acts via ε^2.) A similar definition applies with Y_n replacing Y_n^*. It follows from an examination of the action of I_p on Y_λ that

$$H^1_{\text{str}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y_n) = H^1_{\text{unr}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y_n). \tag{4.4}$$

In the case of Y^* we will use the inequality

$$\#\, H^1_{\text{str}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y^*) \le \#\, H^1_{\text{unr}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y^*). \tag{4.5}$$

We also need the fact that for n sufficiently large the map

$$H^1_{\text{str}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y_n^*) \to H^1_{\text{str}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y^*) \tag{4.6}$$

is injective. One can check this by replacing these groups by the subgroups of $H^1(L, (K/\mathcal{O})(\nu)_{\lambda^n})$ and $H^1(L, (K/\mathcal{O})(\nu))$ which are unramified outside $\mathfrak{p}$ and trivial at $\mathfrak{p}^*$, in a manner similar to the beginning of the proof of Proposition 4.1. The above map is then injective whenever the connecting homomorphism

$$H^0(L_{\mathfrak{p}^*}, (K/\mathcal{O})(\nu)) \to H^1(L_{\mathfrak{p}^*}, (K/\mathcal{O})(\nu)_{\lambda^n})$$

is injective, which holds for sufficiently large n.

Now, by Proposition 1.6,

$$\frac{\#\, H^1_{\text{str}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y_n)}{\#\, H^1_{\text{str}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y_n^*)} = \#\, H^0(\mathbf{Q}_p,\, (Y_n^0)^*)\frac{\#\, H^0\,(\mathbf{Q},\, Y_n)}{\#\, H^0\,(\mathbf{Q},\, Y_n^*)}. \tag{4.7}$$

Also, $H^0\,(\mathbf{Q}, Y_n) = 0$ and a simple calculation shows that

$$\#\, H^0(\mathbf{Q},\, Y_n^*) = \begin{cases} \inf\limits_{\mathfrak{q}}\ \#(\mathcal{O}/1-\nu(\mathfrak{q})) & \text{if } \nu \equiv 1 \bmod \lambda \\ \quad 1 & \text{otherwise} \end{cases}$$

where $\mathfrak{q}$ runs through a set of primes of $\mathcal{O}_L$ prime to $p\,\text{cond}(\nu)$ of density one. This can be checked since $Y^* = \text{Ind}_L^{\mathbf{Q}}(\nu) \underset{\mathcal{O}}{\otimes} K/\mathcal{O}$. So, setting

$$t = \begin{cases} \inf_{\mathfrak{q}}\ \#(\mathcal{O}/(1-\nu(\mathfrak{q}))) & \text{if } \nu \bmod \lambda = 1 \\ 1 & \text{if } \nu \bmod \lambda \ne 1 \end{cases} \tag{4.8}$$

we get

(4.9)

$$\#\, H^1_{\text{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q},\, Y) \le \frac{1}{t}\cdot \prod_{q\in\Sigma} \ell_q \cdot \#\ \text{Hom}\,(\text{Gal}\,(M_\infty/L(\nu))\,,\ (K/\mathcal{O})(\nu))^{\text{Gal}(L(\nu)/L)}$$

where $\ell_q = \#\, H^0(\mathbf{Q}_q,\, Y^*)$ for $q \ne p$, $\ell_p = \lim\limits_{n\to\infty} \#\, H^0(\mathbf{Q}_p,\, (Y_n^0)^*)$. This follows from Proposition 4.1, (4.4)–(4.7) and the elementary estimate

$$\#(H^1_{\text{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y)/H^1_{\text{unr}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y)) \le \prod_{q\in\Sigma-\{p\}} \ell_q, \tag{4.10}$$

which follows from the fact that $\# H^1(\mathbf{Q}_q^{\text{unr}}, Y)^{\text{Gal}(\mathbf{Q}_q^{\text{unr}}/\mathbf{Q}_q)} = \ell_q$.

Our objective is to compute $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$ and the main problem is to estimate $H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, Y)$. By (4.5) this in turn reduces to the problem of estimating
$\# \operatorname{Hom}(\operatorname{Gal}(M_\infty/L(\nu)), (K/\mathcal{O})(\nu))^{\operatorname{Gal}(L(\nu)/L)}$. This order can be computed using the 'main conjecture' established by Rubin using ideas of Kolyvagin. (cf. [Ru2] and especially [Ru4]. In the former reference Rubin assumes that the class number of L is prime to p.) We could now derive the result directly from this by referring to [de Sh, Ch. 3], but we will recall some of the steps here.

Let $w_\mathfrak{f}$ denote the number of roots of unity ζ of L such that $\zeta \equiv 1 \bmod \mathfrak{f}$ ($\mathfrak{f}$ an integral ideal of $\mathcal{O}_L$). We choose an $\mathfrak{f}$ prime to p such that $w_\mathfrak{f} = 1$. Then there is a grossencharacter φ of L satisfying $\varphi((\alpha)) = \alpha$ for $\alpha \equiv 1 \bmod \mathfrak{f}$ (cf. [de Sh, II.1.4]). According to Weil, after fixing an embedding $\bar{\mathbf{Q}} \hookrightarrow \bar{\mathbf{Q}}_p$ we can associate a p-adic character φ_p to φ (cf. [de Sh, II.1.1 (5)]). We choose an embedding corresponding to a prime above $\mathfrak{p}$ and then we find $\varphi_p = \kappa \cdot \chi$ for some χ of finite order and conductor prime to p. Indeed φ_p and κ are both unramified at $\mathfrak{p}^*$ and satisfy $\varphi_p|_{I_\mathfrak{p}} = \kappa|_{I_\mathfrak{p}} = \varepsilon$ where ε is the cyclotomic character and $I_\mathfrak{p}$ is an inertia group at $\mathfrak{p}$. Without altering $\mathfrak{f}$ we can even choose φ so that the order of χ is prime to p. This is by our hypothesis that κ factored through an extension of the form $\mathbf{Z}_p \oplus T$ with T of order prime to p. To see this pick an abelian splitting field of φ_p and κ whose Galois group has the form $G \oplus G'$ with G a pro-p-group and G' of order prime to p. Then we see that $\varphi|_G$ has conductor dividing $\mathfrak{f}\mathfrak{p}^\infty$. Also the only primes which ramify in a $\mathbf{Z}_p$-extension lie above p so our hypothesis on κ ensures that $\kappa|_G$ has conductor dividing $\mathfrak{f}\mathfrak{p}^\infty$. The same is then true of the p-part of χ which therefore has conductor dividing $\mathfrak{f}$. We can therefore adjust φ so that χ has order prime to p as claimed. We will not however choose φ so that χ is 1 as this would require $\mathfrak{f}\mathfrak{p}^\infty$ to be divisible by $\operatorname{cond}\chi$. However we will make the assumption, by altering $\mathfrak{f}$ if necessary, but still keeping $\mathfrak{f}$ prime to p, that both ν and φ_p have conductor dividing $\mathfrak{f}\mathfrak{p}^\infty$. Thus we replace $\mathfrak{f}\mathfrak{p}^\infty$ by $\text{l.c.m.}\{\mathfrak{f}, \operatorname{cond}\nu\}$.

The grossencharacter φ (or more precisely $\varphi \circ N_{F/L}$) is associated to a (unique) elliptic curve E defined over $F = L(\mathfrak{f})$, the ray class field of conductor $\mathfrak{f}$, with complex multiplication by $\mathcal{O}_L$ and isomorphic over $\mathbf{C}$ to $\mathbf{C}/\mathcal{O}_L$ (cf. [de Sh, II. Lemma 1.4]). We may even fix a Weierstrass model of E over $\mathcal{O}_F$ which has good reduction at all primes above $\mathfrak{p}$. For each prime $\mathfrak{P}$ of F above $\mathfrak{p}$ we have a formal group $\hat{E}_\mathfrak{P}$, and this is a relative Lubin-Tate group with respect to $F_\mathfrak{P}$ over $L_\mathfrak{p}$ (cf. [de Sh, Ch. II, §1.10]). We let $\lambda = \lambda_{\hat{E}_\mathfrak{P}}$ be the logarithm of this formal group.

Let U_∞ be the product of the principal local units at the primes above $\mathfrak{p}$ of $L(\mathfrak{f}\mathfrak{p}^\infty)$; i.e.,

$$U_\infty = \prod_{\mathfrak{P}|\mathfrak{p}} U_{\infty,\mathfrak{P}} \qquad \text{where} \qquad U_{\infty,\mathfrak{P}} = \varprojlim U_{n,\mathfrak{P}},$$

each $U_{n,\mathfrak{P}}$ being the principal local units in $L(\mathfrak{f}\mathfrak{p}^n)_{\mathfrak{P}}$. (Note that the primes of $L(\mathfrak{f})$ above $\mathfrak{p}$ are totally ramified in $L(\mathfrak{f}\mathfrak{p}^\infty)$ so we still call them $\{\mathfrak{P}\}$.) We wish to define certain homomorphisms δ_k on U_∞. These were first introduced in [CW] in the case where the local field $F_{\mathfrak{P}}$ is $\mathbf{Q}_p$.

Assume for the moment that $F_{\mathfrak{P}}$ is $\mathbf{Q}_p$. In this case $\hat{E}_{\mathfrak{P}}$ is isomorphic to the Lubin-Tate group associated to $\pi x + x^p$ where $\pi = \varphi(\mathfrak{p})$. Then letting ω_n be nontrivial roots of $[\pi^n](x) = 0$ chosen so that $[\pi](\omega_n) = \omega_{n-1}$, it was shown in [CW] that to each element $u = \varprojlim u_n \in U_{\infty,\mathfrak{P}}$ there corresponded a unique power series $f_u(T) \in \mathbf{Z}_p[\![T]\!]^\times$ such that $f_u(\omega_n) = u_n$ for $n \geq 1$. The definition of $\delta_{k,\mathfrak{P}}$ $(k \geq 1)$ in this case was then

$$\delta_{k,\mathfrak{P}}(u) = \left(\frac{1}{\lambda'(T)}\frac{d}{dT}\right)^k \log f_u(T)\Big|_{T=0}.$$

It is easy to see that $\delta_{k,\mathfrak{P}}$ gives a homomorphism: $U_\infty \to U_{\infty,\mathfrak{P}} \to \mathcal{O}_\mathfrak{p}$ satisfying $\delta_{k,\mathfrak{P}}(\varepsilon^\sigma) = \theta(\sigma)^k \delta_{k,\mathfrak{P}}(\varepsilon)$ where $\theta : \operatorname{Gal}\left(\overline{F}/F\right) \to \mathcal{O}_\mathfrak{p}^\times$ is the character giving the action on $E[\mathfrak{p}^\infty]$.

The construction of the power series in [CW] does not extend to the case where the formal group has height > 1 or to the case where it is defined over an extension of $\mathbf{Q}_p$. A more natural approach was developed by Coleman [Co] which works in general. (See also [Iw1].) The corresponding generalizations of δ_k were given in somewhat greater generality in [Ru3] and then in full generality by de Shalit [de Sh]. We now summarize these results, thus returning to the general case where $F_{\mathfrak{P}}$ is not assumed to be $\mathbf{Q}_p$.

To an element $u = \varprojlim u_n \in U_\infty$ we can associate a power series $f_{u,\mathfrak{P}}(T) \in \mathcal{O}_{\mathfrak{P}}[[T]]^\times$ where $\mathcal{O}_{\mathfrak{P}}$ is the ring of integers of $F_{\mathfrak{P}}$; see [de Sh, Ch. II §4.5]. (More precisely $f_{u,\mathfrak{P}}(T)$ is the $\mathfrak{P}$-component of the power series described there.) For $\mathfrak{P}$ we will choose the prime above $\mathfrak{p}$ corresponding to our chosen embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_p$. This power series satisfies $u_{n,\mathfrak{P}} = (f_{u,\mathfrak{P}})(\omega_n)$ for all $n > 0$, $n \equiv 0(d)$ where $d = [F_{\mathfrak{P}}: L_\mathfrak{p}]$ and $\{\omega_n\}$ is chosen as before as an inverse system of π^n division points of $\hat{E}_{\mathfrak{P}}$. We define a homomorphism $\delta_k : U_\infty \to \mathcal{O}_{\mathfrak{P}}$ by

$$\delta_k(u) := \delta_{k,\mathfrak{P}}(u) = \left(\frac{1}{\lambda'_{\hat{E}_{\mathfrak{P}}}(T)}\frac{d}{dT}\right)^k \log f_{u,\mathfrak{P}}(T)\Big|_{T=0}. \tag{4.11}$$

Then

$$\delta_k(u^\tau) = \theta(\tau)^k \delta_k(u) \qquad \text{for } \tau \in \operatorname{Gal}(\bar{F}/F) \tag{4.12}$$

where θ again denotes the action on $E[\mathfrak{p}^\infty]$. Now $\theta = \varphi_p$ on $\operatorname{Gal}(\bar{F}/F)$. We actually want a homomorphism on U_∞ with a transformation property corresponding to ν on all of $\operatorname{Gal}(\bar{L}/L)$. Observe that $\nu = \varphi_p^2$ on $\operatorname{Gal}(\bar{F}/F)$. Let S

be a set of coset representatives for $\mathrm{Gal}(\bar{L}/L)/\,\mathrm{Gal}\,(\bar{L}/F)$ and define

$$\Phi_2(u) = \sum_{\sigma \in S} \nu^{-1}(\sigma)\delta_2(u^\sigma) \in \mathcal{O}_{\mathfrak{P}}[\nu]. \tag{4.13}$$

Each term is independent of the choice of coset representative by (4.8) and it is easily checked that

$$\Phi_2(u^\sigma) = \nu(\sigma)\Phi_2(u).$$

It takes integral values in $\mathcal{O}_{\mathfrak{P}}[\nu]$. Let $U_\infty(\nu)$ denote the product of the groups of local principal units at the primes above $\mathfrak{p}$ of the field $L(\nu)$ (by which we mean projective limits of local principal units as before). Then Φ_2 factors through $U_\infty(\nu)$ and thus defines a continuous homomorphism

$$\Phi_2 \colon U_\infty(\nu) \to \mathbf{C}_p.$$

Let $\mathcal{C}_\infty$ be the group of projective limits of elliptic units in $L(\nu)$ as defined in [Ru4]. Then we have a crucial theorem of Rubin (cf. [Ru4], [Ru2]), proved using ideas of Kolyvagin:

THEOREM 4.2. *There is an equality of characteristic ideals as* $\Lambda = \mathbf{Z}_p[[\mathrm{Gal}(L(\nu)/L)]]$*-modules:*

$$\mathrm{char}_\Lambda\,(\mathrm{Gal}\,(M_\infty/L(\nu))) = \mathrm{char}_\Lambda(U_\infty(\nu)/\overline{\mathcal{C}}_\infty).$$

Let $\nu_0 = \nu \bmod \lambda$. For any $\mathbf{Z}_p[\mathrm{Gal}(L(\nu_0)/L)]$-module X we write $X^{(\nu_0)}$ for the maximal quotient of $X \underset{\mathbf{Z_p}}{\otimes} \mathcal{O}$ on which the action of $\mathrm{Gal}(L(\nu_0)/L)$ is via the Teichmüller lift of ν_0. Since $\mathrm{Gal}(L(\nu)/L)$ decomposes into a direct product of a pro-p group and a group of order prime to p,

$$\mathrm{Gal}(L(\nu)/L) \simeq \mathrm{Gal}(L(\nu)/L(\nu_0)) \times \mathrm{Gal}(L(\nu_0)/L),$$

we can also consider any $\mathbf{Z}_p[[\mathrm{Gal}(L(\nu)/L)]]$-module also as a $\mathbf{Z}_p[\mathrm{Gal}(L(\nu_0)/L)]$-module. In particular $X^{(\nu_0)}$ is a module over $\mathbf{Z}_p[\mathrm{Gal}(L(\nu_0)/L)]^{(\nu_0)} \simeq \mathcal{O}$. Also $\Lambda^{(\nu_0)} \simeq \mathcal{O}[[T]]$.

Now according to results of Iwasawa ([Iw2, §12], [Ru2, Theorem 5.1]), $U_\infty(\nu)^{(\nu_0)}$ is a free $\Lambda^{(\nu_0)}$-module of rank one. We extend Φ_2 $\mathcal{O}$-linearly to $U_\infty(\nu) \otimes_{\mathbf{Z}_p} \mathcal{O}$ and it then factors through $U_\infty(\nu)^{(\nu_0)}$. Suppose that u is a generator of $U_\infty(\nu)^{(\nu_0)}$ and β an element of $\bar{\mathcal{C}}_\infty^{(\nu_0)}$. Then $f(\gamma-1)u = \beta$ for some $f(T) \in \mathcal{O}[[T]]$ and γ a topological generator of $\mathrm{Gal}\,(L(\nu)/L(\nu_0))$. Computing Φ_2 on both u and β gives

$$f(\nu(\gamma) - 1) = \Phi_2(\beta)/\Phi_2(u). \tag{4.14}$$

Next we let $e(\mathfrak{a})$ be the projective limit of elliptic units in $\varprojlim L^\times_{\mathfrak{f}\mathfrak{p}^n}$ for $\mathfrak{a}$ some ideal prime to $6\mathfrak{f}\mathfrak{p}$ described in [de Sh, Ch. II, §4.9]. Then by the proposition of Chapter II, §2.7 of [de Sh] this is a 12[th] power in $\varprojlim L^\times_{\mathfrak{f}\mathfrak{p}^n}$. We

let $\beta_1 = \beta(\mathfrak{a})^{1/12}$ be the projection of $e(\mathfrak{a})^{1/12}$ to U_∞ and take $\beta = \operatorname{Norm}\beta_1$ where the norm is from $L_{\mathfrak{f}\mathfrak{p}^\infty}$ to $L(\nu)$. A generalization of the calculation in [CW] which may be found in [de Sh, Ch. II, §4.10] shows that

$$\Phi_2(\beta) = \text{(root of unity)}\ \Omega^{-2}\ (N\mathfrak{a} - \nu(\mathfrak{a}))\ L_{\mathfrak{f}}(2, \bar{\nu}) \in \mathcal{O}_{\mathfrak{P}}[\nu] \tag{4.15}$$

where Ω is a basis for the $\mathcal{O}_L$-module of periods of our chosen Weierstrass model of $E_{/F}$. (Recall that this was chosen to have good reduction at primes above $\mathfrak{p}$. The periods are those of the standard Neron differential.) Also ν here should be interpreted as the grossencharacter whose associated p-adic character, via the chosen embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_p$, is ν, and $\bar{\nu}$ is the complex conjugate of ν.

The only restrictions we have placed on $\mathfrak{f}$ are that (i) $\mathfrak{f}$ is prime to $\mathfrak{p}$; (ii) $w_{\mathfrak{f}} = 1$; and (iii) $\operatorname{cond}\nu \mid \mathfrak{f}\mathfrak{p}^\infty$. Now let $\mathfrak{f}_0\mathfrak{p}^\infty$ be the conductor of ν with $\mathfrak{f}_0$ prime to p. We show now that we can choose $\mathfrak{f}$ such that $L_{\mathfrak{f}}(2, \bar{\nu})/L_{\mathfrak{f}_0}(2, \bar{\nu})$ is a p-adic unit unless $\nu_0 = 1$ in which case we can choose it to be t as defined in (4.4). We can clearly choose $L_{\mathfrak{f}}(2, \bar{\nu})/L_{\mathfrak{f}_0}(2, \bar{\nu})$ to be a unit if $\nu_0 \neq 1$, as $\bar{\nu}(\mathfrak{q})\nu(\mathfrak{q}) = \operatorname{Norm}\mathfrak{q}^2$ for any ideal $\mathfrak{q}$ prime to $\mathfrak{f}_0\mathfrak{p}$. Note that if $\nu_0 = 1$ then also $p = 3$. Also if $\nu_0 = 1$ then we see that

$$\inf_{\mathfrak{q}} \# \left\{ \mathcal{O}/\{L_{\mathfrak{f}_0\mathfrak{q}}(2, \bar{\nu})/L_{\mathfrak{f}_0}(2, \bar{\nu})\} \right\} = t$$

since $\bar{\nu}\varepsilon^{-2} = \nu^{-1}$.

We can compute $\Phi_2(u)$ by choosing a special local unit and showing that $\Phi_2(u)$ is a p-adic unit, but it is sufficient for us to know that it is integral. Then since $\operatorname{Gal}(M_\infty/L(\nu))$ has no finite Λ-submodule (by a result of Greenberg; see [Gre2, end of §4]) we deduce from Theorem 4.2, (4.14) and (4.15) that

$$\# \operatorname{Hom}(\operatorname{Gal}(M_\infty/L(\nu)), (K/\mathcal{O})(\nu))^{\operatorname{Gal}(L(\nu)/L)}$$
$$\leq \begin{cases} \#\mathcal{O}/\Omega^{-2}L_{\mathfrak{f}_0}(2, \bar{\nu}) & \text{if } \nu_0 \neq 1 \\ (\#\mathcal{O}/\Omega^{-2}L_{\mathfrak{f}_0}(2, \bar{\nu})) \cdot t & \text{if } \nu_0 = 1. \end{cases}$$

Combining this with (4.9) gives:

$$\#\ H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q},\ Y)\ \leq\ \#\ \left(\mathcal{O}/\Omega^{-2}L_{\mathfrak{f}_0}(2, \bar{\nu})\right) \cdot \prod_{q \in \Sigma} \ell_q$$

where $\ell_q = \#\ H^0(\mathbf{Q}_q,\ Y^*)$ (for $q \neq p$), $\ell_p = \#\ H^0\,(\mathbf{Q}_p,\ (Y^0)^*)$.

Since $V \simeq Y \oplus (K/\mathcal{O})(\psi) \oplus K/\mathcal{O}$ we need also a formula for

$$\#\ \ker\left\{ H^1(\mathbf{Q}_\Sigma/\mathbf{Q}, (K/\mathcal{O})(\psi) \oplus K/\mathcal{O}) \to H^1(\mathbf{Q}_p^{\mathrm{unr}}, (K/\mathcal{O})(\psi) \oplus K/\mathcal{O}) \right\}.$$

This is easily computed to be

$$\#(\mathcal{O}/h_L) \cdot \prod_{q \in \Sigma - \{p\}} \ell_q \tag{4.16}$$

where $\ell_q = \#H^0(\mathbf{Q}_q, ((K/\mathcal{O})(\psi) \oplus K/\mathcal{O})^*)$ and h_L is the class number of $\mathcal{O}_L$. Combining these gives:

PROPOSITION 4.3.

$$\#H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V) \leq \#(\mathcal{O}/\Omega^{-2}L_{\mathfrak{f}_0}(2,\overline{\nu})) \cdot \#(\mathcal{O}/h_L) \cdot \prod_{q\in\Sigma} \ell_q$$

where $\ell_q = \#H^0(\mathbf{Q}_q, V^*)$ *(for* $q \neq p$*)*, $\ell_p = \#H^0(\mathbf{Q}_p, (Y^0)^*)$.

2. Calculation of η

We need to calculate explicitly the invariants $\eta_{\mathcal{D},f}$ introduced in Chapter 2, §3 in a special case. Let ρ_0 be an irreducible representation as in (1.1). Suppose that f is a newform of weight 2 and level N, λ a prime of $\mathcal{O}_f$ above p and $\rho_{f,\lambda}$ a deformation of ρ_0. Let $\mathfrak{m}$ be the kernel of the homomorphism $\mathbf{T}_1(N) \to \mathcal{O}_f/\lambda$ arising from f. We write T for $\mathbf{T}_1(N)_\mathfrak{m} \underset{W(k_\mathfrak{m})}{\otimes} \mathcal{O}$, where $\mathcal{O} = \mathcal{O}_{f,\lambda}$ and $k_\mathfrak{m}$ is the residue field of $\mathfrak{m}$. Assume that $p \nmid N$. We assume here that k is the residue field of $\mathcal{O}$ and that it is chosen to contain $k_\mathfrak{m}$. Then by Corollary 1 of Theorem 2.1, $\mathbf{T}_1(N)_\mathfrak{m}$ is Gorenstein and it follows that T is also a Gorenstein $\mathcal{O}$-algebra (see the discussion following (2.42)). So we can use perfect pairings (the second one T-bilinear)

$$\mathcal{O} \times \mathcal{O} \to \mathcal{O}, \qquad \langle\, , \rangle : T \times T \to \mathcal{O}$$

to define an invariant η of T. If $\pi : T \to \mathcal{O}$ is the natural map, we set $(\eta) = (\hat{\pi}(1))$ where $\hat{\pi}$ is the adjoint of π with respect to the pairings. It is well-defined as an ideal of T, depending only on π. Furthermore, as we noted in Chapter 2, §3, $\pi(\eta) = \langle \eta, \eta \rangle$ up to a unit in $\mathcal{O}$ and as noted in the appendix $\eta = \operatorname{Ann}\mathfrak{p} = T[\mathfrak{p}]$ where $\mathfrak{p} = \ker \pi$. We now give an explicit formula for η developed by Hida (cf. [Hi2] for a survey of his earlier results) by interpreting $\langle\, , \rangle$ in terms of the cup product pairing on the cohomology of $X_1(N)$, and then in terms of the Petersson inner product of f with itself. The following account (which does not require the CM hypothesis) is adapted from [Hi2] and we refer there for more details.

Let

$$(4.17) \qquad (\, , \,) : H^1(X_1(N), \mathcal{O}_f) \times H^1(X_1(N), \mathcal{O}_f) \to \mathcal{O}_f$$

be the cup product pairing with $\mathcal{O}_f$ as coefficients. (We sometimes drop the $\mathbf{C}$ from $X_1(N)_{/\mathbf{C}}$ or $J_1(N)_{/\mathbf{C}}$ if the context makes it clear that we are referring to the complex manifolds.) In particular $(t_* x, y) = (x, t^* y)$ for all x, y and for each standard Hecke correspondence t. We use the action of t on $H^1(X_1(N), \mathcal{O}_f)$ given by $x \mapsto t^* x$ and simply write tx for $t^* x$. This is the same

as the action induced by $t_* \in \mathbf{T}_1(N)$ on $H^1(J_1(N), \mathcal{O}_f) \simeq H^1(X_1(N), \mathcal{O}_f)$. Let $\mathfrak{p}_f$ be the minimal prime of $\mathbf{T}_1(N) \otimes \mathcal{O}_f$ associated to f (i.e., the kernel of $\mathbf{T}_1(N) \otimes \mathcal{O}_f \rightarrow \mathcal{O}_f$ given by $t_l \otimes \beta \mapsto \beta\, c_t(f)$ where $tf = c_t(f)f$), and let

$$L_f = H^1(X_1(N),\, \mathcal{O}_f)\, [\mathfrak{p}_f].$$

If $f = \Sigma a_n q^n$ let $f^\rho = \Sigma \bar{a}_n q^n$. Then f^ρ is again a newform and we define L_{f^ρ} by replacing f by f^ρ in the definition of L_f. (Note here that $\mathcal{O}_f = \mathcal{O}_{f^\rho}$ as these rings are the integers of fields which are either totally real or CM by a result of Shimura. Actually this is not essential as we could replace $\mathcal{O}_f$ by any ring of integers containing it.) Then the pairing (,) induces another by restriction

$$(\,,\,):\ L_f \times L_{f^\rho} \rightarrow \mathcal{O}_f. \tag{4.18}$$

Replacing $\mathcal{O}_f$ (and the $\mathcal{O}_f$-modules) by the localization of $\mathcal{O}_f$ at p (if necessary) we can assume that L_f and L_{f^ρ} are free of rank 2 and direct summands as $\mathcal{O}_f$-modules of the respective cohomology groups. Let δ_1, δ_2 be a basis of L_f. Then also $\bar{\delta}_1, \bar{\delta}_2$ is a basis of $L_{f^\rho} = \overline{L_f}$. Here complex conjugation acts on $H^1(X_1(N),\, \mathcal{O}_f)$ via its action on $\mathcal{O}_f$. We can then verify that

$$(\boldsymbol{\delta},\, \bar{\boldsymbol{\delta}}) := \det(\delta_i,\, \bar{\delta}_j)$$

is an element of $\mathcal{O}_f$ (or its localization at p) whose image in $\mathcal{O}_{f,\lambda}$ is given by $\pi(\eta^2)$ (unit). To see this, consider a modified pairing $\langle\ ,\ \rangle$ defined by

$$\langle x, y\rangle = (x, w_\zeta y) \tag{4.19}$$

where w_ζ is defined as in (2.4). Then $\langle tx, y\rangle = \langle x, ty\rangle$ for all x, y and Hecke operators t. Furthermore

$$\det\langle \delta_i, \delta_j\rangle = \det(\delta_i, w_\zeta \delta_j) = c \det(\delta_i, \bar{\delta}_j)$$

for some p-adic unit c (in $\mathcal{O}_f$). This is because $w_\zeta(L_{f^\rho}) = L_f$ and $w_\zeta(L_f) = L_{f^\rho}$. (One can check this, for example, using the explicit bases described below.) Moreover, by Theorem 2.1,

$$\begin{aligned} H^1(X_1(N),\, \mathbf{Z})\ \otimes_{\mathbf{T}_1(N)} \mathbf{T}_1(N)_\mathfrak{m} \ &\simeq\ \mathbf{T}_1(N)^2_\mathfrak{m}, \\ H^1(X_1(N),\, \mathcal{O}_f)\ \otimes_{\mathbf{T}_1(N)\otimes \mathcal{O}_f} T \ &\simeq\ T^2. \end{aligned}$$

Thus (4.18) can be viewed (after tensoring with $\mathcal{O}_{f,\lambda}$ and modifying it as in (4.19)) as a perfect pairing of T-modules and so this serves to compute $\pi(\eta^2)$ as explained earlier (the square coming from the fact that we have a rank 2 module).

To give a more useful expression for $(\boldsymbol{\delta},\, \bar{\boldsymbol{\delta}})$ we observe that f and $\overline{f^\rho}$ can be viewed as elements of $H^1\,(X_1(N), \mathbf{C}) \simeq H^1_{\rm DR}(X_1(N), \mathbf{C})$ via $f \mapsto f(z)dz$, $\overline{f^\rho} \mapsto \overline{f^\rho} d\bar{z}$. Then $\{f,\, \overline{f^\rho}\}$ form a basis for $L_f \otimes_{\mathcal{O}_f} \mathbf{C}$. Similarly $\{\bar{f},\, f^\rho\}$ form a basis

for $L_{f^\rho} \otimes_{\mathcal{O}_f} \mathbf{C}$. Define the vectors $\boldsymbol{\omega}_1 = (f, \overline{f^\rho})$, $\boldsymbol{\omega}_2 = (\bar{f}, f^\rho)$ and write $\boldsymbol{\omega}_1 = C\boldsymbol{\delta}$ and $\boldsymbol{\omega}_2 = \bar{C}\bar{\boldsymbol{\delta}}$ with $C \in M_2(\mathbf{C})$. Then writing $f_1 = f$, $f_2 = \overline{f^\rho}$ we set

$$(\boldsymbol{\omega}, \bar{\boldsymbol{\omega}}) := \det((f_i, \overline{f_j})) = (\boldsymbol{\delta}, \bar{\boldsymbol{\delta}}) \det(C\bar{C}).$$

Now $(\boldsymbol{\omega}, \bar{\boldsymbol{\omega}})$ is given explicitly in terms of the (non-normalized) Petersson inner product $\langle , \rangle$:

$$(\boldsymbol{\omega}, \bar{\boldsymbol{\omega}}) = -4\langle f, f\rangle^2$$

where $\langle f, f\rangle = \int_{\mathfrak{H}/\Gamma_1(N)} f\bar{f} dx dy$.

To compute $\det(C)$ we consider integrals over classes in $H_1(X_1(N), \mathcal{O}_f)$. By Poincaré duality there exist classes c_1, c_2 in $H_1(X_1(N), \mathcal{O}_f)$ such that $\det(\int_{c_j} \delta_i)$ is a unit in $\mathcal{O}_f$. Hence $\det C$ generates the same $\mathcal{O}_f$-module as is generated by $\left\{\det\left(\int_{c_j} f_i\right)\right\}$ for all such choices of classes (c_1, c_2) and with $\{f_1, f_2\} = \{f, \overline{f^\rho}\}$. Letting u_f be a generator of the $\mathcal{O}_f$-module $\left\{\det\left(\int_{c_j} f_i\right)\right\}$ we have the following formula of Hida:

PROPOSITION 4.4. *$\pi(\eta^2) = \langle f, f\rangle^2/u_f\bar{u}_f \times$ (unit in $\mathcal{O}_{f,\lambda}$).*

Now we restrict to the case where $\rho_0 = \operatorname{Ind}_L^{\mathbf{Q}} \kappa_0$ for some imaginary quadratic field L which is unramified at p and some $k^\times$-valued character κ_0 of $\operatorname{Gal}(\bar{L}/L)$. We assume that ρ_0 is irreducible, i.e., that $\kappa_0 \neq \kappa_{0,\sigma}$ where $\kappa_{0,\sigma}(\delta) = \kappa_0(\sigma^{-1}\delta\sigma)$ for any σ representing the nontrivial coset of $\operatorname{Gal}(\bar{L}/\mathbf{Q})/\operatorname{Gal}(\bar{L}/L)$. In addition we wish to assume that ρ_0 is ordinary and $\det \rho_0|_{I_p} = \omega$. In particular p splits in L. These conditions imply that, if $\mathfrak{p}$ is a prime of L above p, $\kappa_0(\alpha) \equiv \alpha^{-1} \bmod \mathfrak{p}$ on $U_\mathfrak{p}$ after possible replacement of κ_0 by $\kappa_{0,\sigma}$. Here the $U_\mathfrak{p}$ are the units of $L_\mathfrak{p}$ and since κ_0 is a character, the restriction of κ_0 to an inertia group $I_\mathfrak{p}$ induces a homomorphism on $U_\mathfrak{p}$. We assume now that $\mathfrak{p}$ is fixed and κ_0 chosen to satisfy this congruence. Our choice of κ_0 will imply that the grossencharacter introduced below has conductor prime to p.

We choose a (primitive) grossencharacter φ on L together with an embedding $\overline{\mathbf{Q}} \hookrightarrow \overline{\mathbf{Q}}_p$ corresponding to the prime $\mathfrak{p}$ above p such that the induced p-adic character φ_p has the properties:

(i) $\varphi_p \bmod \bar{\mathfrak{p}} = \kappa_0$ ($\bar{\mathfrak{p}}$ = maximal ideal of $\overline{\mathbf{Q}}_p$).

(ii) φ_p factors through an abelian extension isomorphic to $\mathbf{Z}_p \oplus T$ with T of finite order prime to p.

(iii) $\varphi((\alpha)) = \alpha$ for $\alpha \equiv 1(\mathfrak{f})$ for some integral ideal $\mathfrak{f}$ prime to p.

To obtain φ it is necessary first to define φ_p. Let M_∞ denote the maximal abelian extension of L which is unramified outside $\mathfrak{p}$. Let $\theta\colon \operatorname{Gal}(M_\infty/L) \to \overline{\mathbf{Q}}_p^\times$ be any character which factors through a $\mathbf{Z}_p$-extension and induces the

homomorphism $\alpha \mapsto \alpha^{-1}$ on $U_{\mathfrak{p},1} \hookrightarrow \mathrm{Gal}(M_\infty/L)$ where $U_{\mathfrak{p},1} = \{u \in U_{\mathfrak{p}}: u \equiv 1(\mathfrak{p})\}$. Then set $\varphi_p = \kappa_0\theta$, and pick a grossencharacter φ such that $(\varphi)_p = \varphi_p$. Note that our choice of φ here is not necessarily intended to be the same as the choice of grossencharacter in Section 1.

Now let $\mathfrak{f}_\varphi$ be the conductor of φ and let F be the ray class field of conductor $\mathfrak{f}_\varphi\bar{\mathfrak{f}}_\varphi$. Then over F there is an elliptic curve, unique up to isomorphism, with complex multiplication by $\mathcal{O}_L$ and period lattice free, of rank one over $\mathcal{O}_L$ and with associated grossencharacter $\varphi \circ N_{F/L}$. The curve $E_{/F}$ is the extension of scalars of a unique elliptic curve $E_{/F^+}$ where F^+ is the real subfield of F of index 2. (See [Sh1, (5.4.3)].) Over F^+ this elliptic curve has only the p-power isogenies of the form $\pm p^m$ for $m \in \mathbf{Z}$. To see this observe that F is unramified at p and ρ_0 is ordinary so that the only isogenies of degree p over F are the ones that correspond to division by $\ker \mathfrak{p}$ and $\ker \mathfrak{p}'$ where $\mathfrak{p}\mathfrak{p}' = (p)$ in L. Over F^+ these two subgroups are interchanged by complex conjugation, which gives the assertion. We let $E/_{\mathcal{O}_{F^+,(p)}}$ denote a Weierstrass model over $\mathcal{O}_{F^+,(p)}$, the localization of $\mathcal{O}_{F^+}$ at p, with good reduction at the primes above p. Let ω_E be a Neron differential of $E_{/\mathcal{O}_{F^+,(p)}}$. Let Ω be a basis for the $\mathcal{O}_L$-module of periods of ω_E. Then $\overline{\Omega} = u \cdot \Omega$ for some p-adic unit in $F^\times$.

According to a theorem of Hecke, φ is associated to a cusp form f_φ in such a way that the L-series $L(s, \varphi)$ and $L(s, f_\varphi)$ are equal (cf. [Sh4, Lemma 3]). Moreover since φ was assumed primitive, $f = f_\varphi$ is a newform. Thus the integer $N = \mathrm{cond}\, f = |\Delta_{L/\mathbf{Q}}|\,\mathrm{Norm}_{L/\mathbf{Q}}(\mathrm{cond}\,\varphi)$ is prime to p and there is a homomorphism

$$\psi_f:\ \mathbf{T}_1(N) \twoheadrightarrow R_f \subset \mathcal{O}_f \subset \mathcal{O}_\varphi$$

satisfying $\psi_f(T_l) = \varphi(\mathfrak{c}) + \varphi(\bar{\mathfrak{c}})$ if $l = \mathfrak{c}\bar{\mathfrak{c}}$ in L, $(l \nmid N)$ and $\psi_f(T_l) = 0$ if l is inert in L $(l \nmid N)$. Also $\psi_f(l\langle l\rangle) = \varphi((l))\psi(l)$ where ψ is the quadratic character associated to L. Using the embedding of $\bar{\mathbf{Q}}$ in $\bar{\mathbf{Q}}_p$ chosen above we get a prime λ of $\mathcal{O}_f$ above p, a maximal ideal $\mathfrak{m}$ of $\mathbf{T}_1(N)$ and a homomorphism $\mathbf{T}_1(N)_\mathfrak{m} \to \mathcal{O}_{f,\lambda}$ such that the associated representation $\rho_{f,\lambda}$ reduces to $\rho_0 \bmod \lambda$.

Let $\mathfrak{p}_0 = \ker \psi_f: \mathbf{T}_1(N) \to \mathcal{O}_f$ and let

$$A_f = J_1(N)/\mathfrak{p}_0 J_1(N)$$

be the abelian variety associated to f by Shimura. Over F^+ there is an isogeny

$$A_{f/F^+} \sim (E_{/F^+})^d$$

where $d = [\mathcal{O}_f: \mathbf{Z}]$ (see [Sh4, Th. 1]). To see this one checks that the p-adic Galois representations associated to the Tate modules on each side are equivalent to $(\mathrm{Ind}_F^{F^+} \varphi_p) \otimes_{\mathbf{Z}_p} K_{f,p}$ where $K_{f,p} = \mathcal{O}_f \otimes \mathbf{Q}_p$ and where $\varphi_p: \mathrm{Gal}(\overline{F}/F) \to \mathbf{Z}_p^\times$ is the p-adic character associated to φ and restricted to F. (One compares trace(Frob ℓ) in the two representations for $\ell \nmid Np$ and ℓ split completely in F^+; cf. the discussion after Theorem 2.1 for the representation on A_f.)

Now pick a nonconstant map

$$\pi: X_1(N)_{/F^+} \to E_{/F^+}$$

which factors through A_{f/F^+}. Let M be the composite of F^+ and the normal closure of K_f viewed in $\mathbf{C}$. Let ω_E be a Neron differential of $E_{/\mathcal{O}_{F^+,(p)}}$. Extending scalars to M we can write

$$\pi^*\omega_E = \sum_{\sigma\in\mathrm{Hom}(K_f,\mathbf{C})} a_\sigma \omega_{f^\sigma}, \qquad a_\sigma \in M$$

where $\omega_{f^\sigma} = \sum_{n=1}^{\infty} a_n(f^\sigma) q^n \frac{dq}{q}$ for each σ. By suitably choosing π we can assume that $a_{\mathrm{id}} \neq 0$. Then there exist $\lambda_i \in \mathcal{O}_M$ and $t_i \in \mathbf{T}_1(N)$ such that

$$\sum \lambda_i t_i \pi^* \omega_E = c_1 \omega_f \qquad \text{for some} \qquad c_1 \in M.$$

We consider the map

$$\pi': H_1(X_1(N)_{/\mathbf{C}}, \mathbf{Z}) \otimes \mathcal{O}_{M,(p)} \to H_1(E_{/\mathbf{C}}, \mathbf{Z}) \otimes \mathcal{O}_{M,(p)} \tag{4.20}$$

given by $\pi' = \sum \lambda_i(\pi \circ t_i)$. Even if π' is not surjective we claim that the image of π' always has the form $H_1(E_{/\mathbf{C}}, \mathbf{Z}) \otimes a\mathcal{O}_{M,(p)}$ for some $a \in \mathcal{O}_M$. This is because tensored with $\mathbf{Z}_p$ π' can be viewed as a $\mathrm{Gal}(\overline{\mathbf{Q}}/F^+)$-equivariant map of p-adic Tate-modules, and the only p-power isogenies on $E_{/F^+}$ have the form $\pm p^m$ for some $m \in \mathbf{Z}$. It follows that we can factor π' as $(1 \otimes a) \circ \alpha$ for some other surjective α

$$\alpha: H_1(X_1(N)_{/\mathbf{C}}, \mathbf{Z}) \otimes \mathcal{O}_M \to H^1(E_{/\mathbf{C}}, \mathbf{Z}) \otimes \mathcal{O}_M,$$

now allowing a to be in $\mathcal{O}_{M,(p)}$. Now define α^* on $\Omega^1_{E/\mathbf{C}}$ by $\alpha^* = \sum a^{-1}\lambda_i t_i \circ \pi^*$ where $\pi^*: \Omega^1_{E/\mathbf{C}} \to \Omega^1_{J_1(N)/\mathbf{C}}$ is the map induced by π and t_i has the usual action on $\Omega^1_{J_1(N)/\mathbf{C}}$. Then $\alpha^*(\omega_E) = c\omega_f$ for some $c \in M$ and

$$\int_\gamma \alpha^*(\omega_E) = \int_{\alpha(\gamma)} \omega_E \tag{4.21}$$

for any class $\gamma \in H_1(X_1(N)_{/\mathbf{C}}, \mathcal{O}_M)$. We note that α (on homology as in (4.20)) also comes from a map of abelian varieties $\alpha: J_1(N)_{/F^+} \otimes_{\mathbf{Z}} \mathcal{O}_M \to E_{/F^+} \otimes_{\mathbf{Z}} \mathcal{O}_M$ although we have not used this to define α^*.

We claim now that $c \in \mathcal{O}_{M,(p)}$. We can compute $\alpha^*(\omega_E)$ by considering $\alpha^*(\omega_E \otimes 1) = \sum t_i \pi^* \otimes a^{-1}\lambda_i$ on $\Omega^1_{E/F^+} \otimes \mathcal{O}_M$ and then mapping the image in $\Omega^1_{J_1(N)/F^+} \otimes \mathcal{O}_M$ to $\Omega^1_{J_1(N)/F^+} \otimes_{\mathcal{O}_{F^+}} \mathcal{O}_M = \Omega^1_{J_1(N)/M}$. Now let us write $\mathcal{O}_1$ for $\mathcal{O}_{F^+,(p)}$. Then there are isomorphisms

$$\Omega^1_{J_1(N)/\mathcal{O}_1 \otimes \mathcal{O}_2} \overset{s_1}{\underset{\sim}{\longrightarrow}} \mathrm{Hom}(\mathcal{O}_M, \Omega^1_{J_1(N)/\mathcal{O}_1}) \overset{s_2}{\underset{\sim}{\longrightarrow}} \Omega^1_{J_1(N)/\mathcal{O}_1} \otimes \delta^{-1}$$

where δ is the different of $M/\mathbf{Q}$. The first isomorphism can be described as follows. Let $e(\gamma): J_1(N) \to J_1(N) \otimes \mathcal{O}_M$ for $\gamma \in \mathcal{O}_M$ be the map $x \mapsto x \otimes \gamma$. Then $t_1(\omega)(\gamma) = e(\gamma)^*\omega$. Similar identifications occur for E in place of $J_1(N)$. So to check that $\alpha^*(\omega_E \otimes 1) \in \Omega^1_{J_1(N)/\mathcal{O}_1} \otimes \mathcal{O}_M$ it is enough to observe that by its construction α comes from a homomorphism $J_1(N)_{/\mathcal{O}_1} \otimes \mathcal{O}_M \to E_{/\mathcal{O}_1} \otimes \mathcal{O}_M$. It follows that we can compare the periods of f and of ω_E.

For f^ρ we use the fact that $\overline{\int_\gamma f^\rho dz} = \int_{\gamma^c} f\, dz$ where c is the $\mathcal{O}_M$-linear map on homology coming from complex conjugation on the curve. We deduce:

PROPOSITION 4.5. $u_f = \frac{1}{4\pi^2}\Omega^2.(1/(p\text{-}adic\ integer))$.

We now give an expression for $\langle f_\varphi, f_\varphi \rangle$ in terms of the L-function of φ. This was first observed by Shimura [Sh2] although the precise form we want was given by Hida.

PROPOSITION 4.6.

$$\langle f_\varphi, f_\varphi \rangle = \frac{1}{16\pi^3}\, N^2 \left\{ \prod_{\substack{q|N \\ q\notin S_\varphi}} \left(1 - \frac{1}{q}\right) \right\} L_N(2,\varphi^2\bar{\hat{\chi}})\, L_N(1,\psi)$$

where χ is the character of f_φ and $\hat{\chi}$ its restriction to L;

ψ is the quadratic character associated to L;

$L_N(\quad)$ denotes that the Euler factors for primes dividing N have been removed;

S_φ is the set of primes $q \mid N$ such that $q = \mathfrak{q}\mathfrak{q}'$ with $\mathfrak{q} \nmid$ cond φ *and $\mathfrak{q},\mathfrak{q}'$ primes of L, not necessarily distinct.*

Proof. One begins with a formula of Petersson that for an eigenform of weight 2 on $\Gamma_1(N)$ says

$$\langle f, f \rangle = (4\pi)^{-2}\, \Gamma\,(2) \left(\frac{1}{3}\right) \pi\, [\mathrm{SL}_2(\mathbf{Z}) : \Gamma_1(N) \cdot (\pm 1)] \cdot \mathrm{Res}_{s=2}\, D(s, f, f^\rho)$$

where $D(s, f, f^\rho) = \sum_{n=1}^{\infty} |a_n|^2 n^{-s}$ if $f = \sum_{n=1}^{\infty} a_n q^n$ (cf. [Hi3, (5.13)]). One checks that, removing the Euler factors at primes dividing N,

$$D_N(s, f, f^\rho) = L_N(s,\, \varphi^2\bar{\hat{\chi}})\, L_N(s-1,\, \psi)\zeta_{\mathbf{Q},N}(s-1)/\zeta_{\mathbf{Q},N}(2s-2)$$

by using Lemma 1 of [Sh3]. For each Euler factor of f at a $q \mid N$ of the form $(1-\alpha_q\, q^{-s})$ we get also an Euler factor in $D(s, f, f^\rho)$ of the form $(1-\alpha_q\bar{\alpha}_q\, q^{-s})$. When $f = f_\varphi$ this can only happen for a split prime $\mathfrak{q}$ where $\mathfrak{q}'$ divides the conductor of φ but $\mathfrak{q}$ does not, or for a ramified prime $\mathfrak{q}$ which does not divide the conductor of φ. In this case we get a term $(1 - q^{1-s})$ since $|\varphi\,(\mathfrak{q})|^2 = q$.

Putting together the propositions of this section we now have a formula for $\pi(\eta)$ as defined at the beginning of this section. Actually it is more convenient

to give a formula for $\pi(\eta_M)$, an invariant defined in the same way but with $\mathbf{T}_1(M)_{\mathfrak{m}_1} \underset{W(k_{\mathfrak{m}_1})}{\otimes} \mathcal{O}$ replacing $\mathbf{T}_1(N)_{\mathfrak{m}} \underset{W(k_{\mathfrak{m}})}{\otimes} \mathcal{O}$ where $M = pM_0$ with $p \nmid M_0$ and M/N is of the form

$$\prod_{q \in S_\varphi} q \cdot \prod_{\substack{q \nmid N \\ q \mid M_0}} q^2.$$

Here $\mathfrak{m}_1$ is defined by the requirements that $\rho_{\mathfrak{m}_1} = \rho_0$, $U_q \in \mathfrak{m}$ if $q \mid M$ $(q \neq p)$ and there is an embedding (which we fix) $k_{\mathfrak{m}_1} \hookrightarrow k$ over k_0 taking $U_p \to \alpha_p$ where α_p is the unit eigenvalue of $\mathrm{Frob}\, p$ in $\rho_{f,\lambda}$. So if f' is the eigenform obtained from f by 'removing the Euler factors' at $q \mid (M/N)$ $(q \neq p)$ and removing the non-unit Euler factor at p we have $\eta_M = \hat{\pi}(1)$ where $\pi : T_1 = \mathbf{T}_1(M)_{\mathfrak{m}_1} \underset{W(k_{\mathfrak{m}_1})}{\otimes} \mathcal{O} \to \mathcal{O}$ corresponds to f' and the adjoint is taken with respect to perfect pairings of T_1 and $\mathcal{O}$ with themselves as $\mathcal{O}$-modules, the first one assumed T_1-bilinear.

Property (ii) of φ_p ensures that M is as in (2.24) with $\mathcal{D} = (\mathrm{Se}, \Sigma, \mathcal{O}, \phi)$ where Σ is the set of primes dividing M. (Note that S_φ is precisely the set of primes q for which $n_q = 1$ in the notation of Chapter 2, §3.) As in Chapter 2, §3 there is a canonical map

$$R_{\mathcal{D}} \to \mathbf{T}_{\mathcal{D}} \simeq \mathbf{T}_1(M)_{\mathfrak{m}_1} \underset{W(k_{\mathfrak{m}_1})}{\otimes} \mathcal{O} \tag{4.22}$$

which is surjective by the arguments in the proof of Proposition 2.15. Here we are considering a slightly more general situation than that in Chapter 2, §3 as we are allowing ρ_0 to be induced from a character of $\mathbf{Q}(\sqrt{-3})$. In this special case we define $\mathbf{T}_{\mathcal{D}}$ to be $\mathbf{T}_1(M)_{\mathfrak{m}_1} \underset{W(k_{\mathfrak{m}_1})}{\otimes} \mathcal{O}$. The existence of the map in (4.22) is proved as in Chapter 2, §3. For the surjectivity, note that for each $q \mid M$ (with $q \neq p$) U_q is zero in $\mathbf{T}_{\mathcal{D}}$ as $U_q \in \mathfrak{m}_1$ for each such q so that we can apply Remark 2.8. To see that U_p is in the image of $R_{\mathcal{D}}$ we use that it is the eigenvalue of $\mathrm{Frob}\, p$ on the unique unramified quotient which is free of rank one in the representation ρ described after the corollaries to Theorem 2.1 (cf. Theorem 2.1.4 of [Wil]). To verify this one checks that $\mathbf{T}_{\mathcal{D}}$ is reduced or alternatively one can apply the method of Remark 2.11. We deduce that $U_p \in \mathbf{T}_{\mathcal{D}}^{\mathrm{tr}}$, the $W(k_{\mathfrak{m}_1})$-subalgebra of $\mathbf{T}_1(M)_{\mathfrak{m}_1}$ generated by the traces, and it follows then that it is in the image of $R_{\mathcal{D}}$. We also need to give a definition of $\mathbf{T}_{\mathcal{D}}$ where $\mathcal{D} = (\mathrm{ord}, \Sigma, \mathcal{O}, \phi)$ and ρ_0 is induced from a character of $\mathbf{Q}(\sqrt{-3})$. For this we use (2.31).

Now we take

$$M = Np \prod_{q \in S_\varphi} q.$$

The arguments in the proof of Theorem 2.17 show that

$$\pi(\eta_M) \text{ is divisible by } \pi(\eta)(\alpha_p^2 - \langle p\rangle) \cdot \prod_{q \in S_\varphi} (q-1)$$

where α_p is the unit eigenvalue of Frob p in $\rho_{f,\lambda}$. The factor at p is given by remark 2.18 and at q it comes from the argument of Proposition 2.12 but with $H = H' = 1$. Combining this with Propositions 4.4, 4.5, and 4.6, we have that

$$(4.23)\quad \pi(\eta_M) \text{ is divisible by } \Omega^{-2} L_N\left(2,\, \varphi^2\bar{\chi}\right)\ \frac{L_N(1,\psi)}{\pi}\,(\alpha_p^2 - \langle p\rangle) \prod_{q|N} (q-1).$$

We deduce:

THEOREM 4.7. $\#(\mathcal{O}/\pi(\eta_M)) = \# H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$.

Proof. As explained in Chapter 2, §3 it is sufficient to prove the inequality $\#(\mathcal{O}/\pi(\eta_M)) \geq \# H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$ as the opposite one is immediate. For this it suffices to compare (4.23) with Proposition 4.3. Since

$$L_N(2,\bar{\nu}) = L_N(2,\nu) = L_N(2,\, \varphi^2\bar{\chi})$$

(note that the right-hand term is real by Proposition 4.6) it suffices to pair up the Euler factors at q for $q \mid N$ in (4.23) and in the expression for the upper bound of $\# H^1_{\mathrm{Se}}(\mathbf{Q}_\Sigma/\mathbf{Q}, V)$. □

We now deduce the main theorem in the CM case using the method of Theorem 2.17.

THEOREM 4.8. *Suppose that ρ_0 as in (1.1) is an irreducible representation of odd determinant such that $\rho_0 = \mathrm{Ind}_L^{\mathbf{Q}}\,\kappa_0$ for a character κ_0 of an imaginary quadratic extension L of $\mathbf{Q}$ which is unramified at p. Assume also that:*

(i) $\det \rho_0\big|_{I_p} = \omega$;

(ii) *ρ_0 is ordinary.*

Then for every $\mathcal{D} = (\cdot, \Sigma, \mathcal{O}, \phi)$ such that ρ_0 is of type $\mathcal{D}$ with $\cdot =$ Se or ord,

$$R_{\mathcal{D}} \simeq \mathbf{T}_{\mathcal{D}}$$

and $\mathbf{T}_{\mathcal{D}}$ is a complete intersection.

COROLLARY. *For any ρ_0 as in the theorem suppose that*

$$\rho\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathcal{O})$$

is a continuous representation with values in the ring of integers of a local field, unramified outside a finite set of primes, satisfying $\bar{\rho} \simeq \rho_0$ when viewed as representations to $\mathrm{GL}_2(\bar{\mathbf{F}}_p)$. Suppose further that:

(i) $\rho\big|_{D_p}$ *is ordinary;*

(ii) $\det\rho\big|_{I_p} = \chi\varepsilon^{k-1}$ *with* χ *of finite order,* $k \geq 2$.

Then ρ *is associated to a modular form of weight* k.

Chapter 5

In this chapter we prove the main results about elliptic curves and especially show how to remove the hypothesis that the representation associated to the 3-division points should be irreducible.

Application to elliptic curves

The key result used is the following theorem of Langlands and Tunnell, extending earlier results of Hecke in the case where the projective image is dihedral.

THEOREM 5.1 (Langlands–Tunnell). *Suppose that* $\rho\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$ *is a continuous irreducible representation whose image is finite and solvable. Suppose further that* $\det\rho$ *is odd. Then there exists a weight one newform* f *such that* $L(s,f) = L(s,\rho)$ *up to finitely many Euler factors.*

Langlands actually proved in [La] a much more general result without restriction on the determinant or the number field (which in our case is $\mathbf{Q}$). However in the crucial case where the image in $\mathrm{PGL}_2(\mathbf{C})$ is S_4, the result was only obtained with an additional hypothesis. This was subsequently removed by Tunnell in [Tu].

Suppose then that

$$\rho_0\colon \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{F}_3)$$

is an irreducible representation of odd determinant. We now show, using the theorem, that this representation is modular in the sense that over $\bar{\mathbf{F}}_3$, $\rho_0 \approx \rho_{g,\mu} \bmod \mu$ for some pair (g,μ) with g some newform of weight 2 (cf. [Se, §5.3]). There exists a representation

$$i : \mathrm{GL}_2(\mathbf{F}_3) \hookrightarrow \mathrm{GL}_2\left(\mathbf{Z}\left[\sqrt{-2}\right]\right) \subset \mathrm{GL}_2(\mathbf{C}).$$

By composing i with an automorphism of $\mathrm{GL}_2(\mathbf{F}_3)$ if necessary we can assume that i induces the identity on reduction mod $\left(1+\sqrt{-2}\right)$. So if we consider

$i \circ \rho_0 : \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}) \to \mathrm{GL}_2(\mathbf{C})$ we obtain an irreducible representation which is easily seen to be odd and whose image is solvable. Applying the theorem we find a newform f of weight one associated to this representation. Its eigenvalues lie in $\mathbf{Z}\left[\sqrt{-2}\right]$. Now pick a modular form E of weight one such that $E \equiv 1(3)$. For example, we can take $E = 6\,E_{1,\chi}$ where $E_{1,\chi}$ is the Eisenstein series with Mellin transform given by $\zeta(s)\,\zeta(s,\chi)$ for χ the quadratic character associated to $\mathbf{Q}(\sqrt{-3})$. Then $fE \equiv f \bmod 3$ and using the Deligne-Serre lemma ([DS, Lemma 6.11]) we can find an eigenform g' of weight 2 with the same eigenvalues as f modulo a prime μ' above $(1+\sqrt{-2})$. There is a newform g of weight 2 which has the same eigenvalues as g' for almost all T_l's, and we replace (g', μ') by (g, μ) for some prime μ above $(1+\sqrt{-2})$. Then the pair (g, μ) satisfies our requirements for a suitable choice of μ (compatible with μ').

We can apply this to an elliptic curve E defined over $\mathbf{Q}$ by considering $E[3]$. We now show how in studying elliptic curves our restriction to irreducible representations in the deformation theory can be circumvented.

THEOREM 5.2. *All semistable elliptic curves over $\mathbf{Q}$ are modular.*

Proof. Suppose that E is a semistable elliptic curve over $\mathbf{Q}$. Assume first that the representation $\bar{\rho}_{E,3}$ on $E[3]$ is irreducible. Then if $\rho_0 = \bar{\rho}_{E,3}$ restricted to $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}(\sqrt{-3}))$ were not absolutely irreducible, the image of the restriction would be abelian of order prime to 3. As the semistable hypothesis implies that all the inertia groups outside 3 in the splitting field of ρ_0 have order dividing 3 this means that the splitting field of ρ_0 is unramified outside 3. However, $\mathbf{Q}(\sqrt{-3})$ has no nontrivial abelian extensions unramified outside 3 and of order prime to 3. So ρ_0 itself would factor through an abelian extension of $\mathbf{Q}$ and this is a contradiction as ρ_0 is assumed odd and irreducible. So ρ_0 restricted to $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}(\sqrt{-3}))$ is absolutely irreducible and $\rho_{E,3}$ is then modular by Theorem 0.2 (proved at the end of Chapter 3). By Serre's isogeny theorem, E is also modular (in the sense of being a factor of the Jacobian of a modular curve).

So assume now that $\bar{\rho}_{E,3}$ is reducible. Then we claim that the representation $\bar{\rho}_{E,5}$ on the 5-division points is irreducible. This is because $X_0(15)\,(\mathbf{Q})$ has only four rational points besides the cusps and these correspond to non-semistable curves which in any case are modular; cf. [BiKu, pp. 79–80]. If we knew that $\bar{\rho}_{E,5}$ was modular we could now prove the theorem in the same way we did knowing that $\bar{\rho}_{E,3}$ was modular once we observe that $\bar{\rho}_{E,5}$ restricted to $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q}(\sqrt{5}))$ is absolutely irreducible. This irreducibility follows a similar argument to the one for $\bar{\rho}_{E,3}$ since the only nontrivial abelian extension of $\mathbf{Q}(\sqrt{5})$ unramified outside 5 and of order prime to 5 is $\mathbf{Q}(\zeta_5)$ which is abelian over $\mathbf{Q}$. Alternatively, it is enough to check that there are no elliptic curves E for which $\bar{\rho}_{E,5}$ is an induced representation over $\mathbf{Q}(\sqrt{5})$ and E is semistable

at 5. This can be checked in the supersingular case using the description of $\bar{\rho}_{E,5}|_{D_5}$ (in particular it is induced from a character of the unramified quadratic extension of $\mathbf{Q}_5$ whose restriction to inertia is the fundamental character of level 2) and in the ordinary case it is straightforward.

Consider the twisted form $X(\rho)_{/\mathbf{Q}}$ of $X(5)_{/\mathbf{Q}}$ defined as follows. Let $X(5)_{/\mathbf{Q}}$ be the (geometrically disconnected) curve whose non-cuspidal points classify elliptic curves with full level 5 structure and let the twisted curve be defined by the cohomology class (even homomorphism) in

$$H^1(\mathrm{Gal}(L/\mathbf{Q}), \quad \mathrm{Aut}\, X(5)_{/L})$$

given by $\bar{\rho}_{E,5}$: $\mathrm{Gal}(L/\mathbf{Q}) \longrightarrow \mathrm{GL}_2(\mathbf{Z}/5\mathbf{Z}) \subseteq \mathrm{Aut}\, X(5)_{/L}$ where L denotes the splitting field of $\bar{\rho}_{E,5}$. Then E defines a rational point on $X(\rho)_{/\mathbf{Q}}$ and hence also of an irreducible component of it which we denote C. This curve C is smooth as $X(\rho)_{/\bar{\mathbf{Q}}} = X(5)_{/\bar{\mathbf{Q}}}$ is smooth. It has genus zero since the same is true of the irreducible components of $X(5)_{/\bar{\mathbf{Q}}}$.

A rational point on C (necessarily non-cuspidal) corresponds to an elliptic curve E' over $\mathbf{Q}$ with an isomorphism $E'[5] \simeq E[5]$ as Galois modules (cf. [DR, VI, Prop. 3.2]). We claim that we can choose such a point with the two properties that (i) the Galois representation $\bar{\rho}_{E',3}$ is irreducible and (ii) E' (or a quadratic twist) has semistable reduction at 5. The curve E' (or a quadratic twist) will then satisfy all the properties needed to apply Theorem 0.2. (For the primes $q \neq 5$ we just use the fact that E' is semistable at $q \iff \#\bar{\rho}_{E',5}(I_q) \mid 5$.) So E' will be modular and hence so too will $\bar{\rho}_{E',5}$.

To pick a rational point on C satisfying (i) and (ii) we use the Hilbert irreducibility theorem. For, to ensure condition (i) holds, we only have to eliminate the possibility that the image of $\bar{\rho}_{E',3}$ is reducible. But this corresponds to E' being the image of a rational point on an irreducible covering of C of degree 4. Let $\mathbf{Q}(t)$ be the function field of C. We have therefore an irreducible polynomial $f(x,t) \in \mathbf{Q}\,(t)[x]$ of degree > 1 and we need to ensure that for many values t_0 in $\mathbf{Q}$, $f(x,t_0)$ has no rational solution. Hilbert's theorem ensures that there exists a t_1 such that $f(x,t_1)$ is irreducible. Then we pick a prime $p_1 \neq 5$ such that $f(x,t_1)$ has no root mod p_1. (This is easily achieved using the Čebotarev density theorem; cf. [CF, ex. 6.2, p. 362].) So finally we pick any $t_0 \in \mathbf{Q}$ which is p_1-adically close to t_1 and also 5-adically close to the original value of t giving E. This last condition ensures that E' (corresponding to t_0) or a quadratic twist has semistable reduction at 5. To see this, observe that since $j_E \neq 0$, 1728, we can find a family $E(j)$: $y^2 = x^3 - g_2(j)x - g_3(j)$ with rational functions $g_2(j)$, $g_3(j)$ which are finite at j_E and with the j-invariant of $E(j_0)$ equal to j_0 whenever the $g_i(j_0)$ are finite. Then E is given by a quadratic twist of $E(j_E)$ and so after a change of functions of the form $g_2(j) \mapsto u^2 g_2(j)$, $g_3(j) \mapsto u^3 g_3(j)$ with $u \in \mathbf{Q}^\times$ we can assume that $E(j_E) = E$ and that the equation $E(j_E)$ is minimal at 5. Then for $j' \in \mathbf{Q}$ close enough 5-adically to j_E

the equation $E(j')$ is still minimal and semistable at 5, since a criterion for this, for an integral model, is that either $\mathrm{ord}_5(\triangle(E(j'))) = 0$ or $\mathrm{ord}_5(c_4(E(j'))) = 0$. So up to a quadratic twist E' is also semistable. □

This kind of argument can be applied more generally.

THEOREM 5.3. *Suppose that E is an elliptic curve defined over $\mathbf{Q}$ with the following properties:*

(i) *E has good or multiplicative reduction at* 3, 5,

(ii) *For $p = 3,5$ and for any prime $q \equiv -1 \bmod p$ either $\bar{\rho}_{E,p}|_{D_q}$ is reducible over $\bar{\mathbf{F}}_p$ or $\bar{\rho}_{E,p}|_{I_q}$ is irreducible over $\bar{\mathbf{F}}_p$.*

Then E is modular.

Proof. The main point to be checked is that one can carry over condition (ii) to the new curve E'. For this we use that for any odd prime $p \neq q$,

$$\bar{\rho}_{E,p}|_{D_q} \quad \text{is absolutely irreducible and } \bar{\rho}_{E,p}|_{I_q} \text{ is absolutely reducible and } 3 \nmid \#\bar{\rho}_{E,p}(I_q)$$
$$\Updownarrow$$
$$E \text{ acquires good reduction over an abelian 2-power extension of } \mathbf{Q}_q^{\mathrm{unr}} \text{ but not over an abelian extension of } \mathbf{Q}_q.$$

Suppose then that $q \equiv -1(3)$ and that E' does not satisfy condition (ii) at q (for $p = 3$). Then we claim that also $3 \nmid \#\bar{\rho}_{E',3}(I_q)$. For otherwise $\bar{\rho}_{E',3}(I_q)$ has its normalizer in $\mathrm{GL}_2(\mathbf{F}_3)$ contained in a Borel, whence $\bar{\rho}_{E',3}(D_q)$ would be reducible which contradicts our hypothesis. So using the above equivalence we deduce, by passing via $\bar{\rho}_{E',5} \simeq \bar{\rho}_{E,5}$, that E also does not satisfy hypothesis (ii) at $p = 3$.

We also need to ensure that $\bar{\rho}_{E',3}$ is absolutely irreducible over $\mathbf{Q}(\sqrt{-3})$. This we can do by observing that the property that the image of $\bar{\rho}_{E',3}$ lies in the Sylow 2-subgroup of $\mathrm{GL}_2(\mathbf{F}_3)$ implies that E' is the image of a rational point on a certain irreducible covering of C of nontrivial degree. We can then argue in the same way we did in the previous theorem to eliminate the possibility that $\bar{\rho}_{E',3}$ was reducible, this time using two separate coverings to ensure that the image of $\bar{\rho}_{E',3}$ is neither reducible nor contained in a Sylow 2-subgroup.

Finally one also has to show that if both $\bar{\rho}_{E,5}$ is reducible and $\bar{\rho}_{E,3}$ is induced from a character of $\mathbf{Q}(\sqrt{-3})$ then E is modular. (The case where both were reducible has already been considered.) Taylor has pointed out that curves satisfying both these conditions are classified by the non-cuspidal rational points on a modular curve isomorphic to $X_0(45)/W_9$, and this is an elliptic curve isogenous to $X_0(15)$ with rank zero over $\mathbf{Q}$. The non-cuspidal rational points correspond to modular elliptic curves of conductor 338. □

Appendix

Gorenstein rings and local complete intersections

Proposition 1. *Suppose that $\mathcal{O}$ is a complete discrete valuation ring and that $\varphi: S \to T$ is a surjective local $\mathcal{O}$-algebra homomorphism between complete local Noetherian $\mathcal{O}$-algebras. Suppose further that $\mathfrak{p}_T$ is a prime ideal of T such that $T/\mathfrak{p}_T \xrightarrow{\sim} \mathcal{O}$ and let $\mathfrak{p}_S = \varphi^{-1}(\mathfrak{p}_T)$. Assume that*

(i) *$T \simeq \mathcal{O}[\![x_1, \dots, x_r]\!]/(f_1, \dots, f_{r-u})$ where r is the size of a minimal set of $\mathcal{O}$-generators of $\mathfrak{p}_T/\mathfrak{p}_T^2$,*

(ii) *φ induces an isomorphism $\mathfrak{p}_S/\mathfrak{p}_S^2 \xrightarrow{\sim} \mathfrak{p}_T/\mathfrak{p}_T^2$ and that these are finitely generated $\mathcal{O}$-modules whose free part has rank u.*

Then φ is an isomorphism.

Proof. First we consider the case where $u = 0$. We may assume that the generators $x_1, \dots, x_r$ lie in $\mathfrak{p}_T$ by subtracting their residues in $T/\mathfrak{p}_T \xrightarrow{\sim} \mathcal{O}$. By (ii) we may also write

$$S \simeq \mathcal{O}[\![x_1, \dots, x_r]\!]/(g_1, \dots, g_s)$$

with $s \geq r$ (by allowing repetitions if necessary) and $\mathfrak{p}_S$ generated by the images of $\{x_1, \dots, x_r\}$. Let $\mathfrak{p} = (x_1, \dots, x_r)$ in $\mathcal{O}[\![x_1, \dots, x_r]\!]$. Writing $f_i \equiv \Sigma a_{ij} x_j \bmod \mathfrak{p}^2$ with $a_{ij} \in \mathcal{O}$, we see that the Fitting ideal as an $\mathcal{O}$-module of $\mathfrak{p}_T/\mathfrak{p}_T^2$ is given by

$$F_{\mathcal{O}}(\mathfrak{p}_T/\mathfrak{p}_T^2) = \det(a_{ij}) \in \mathcal{O}$$

and that this is nonzero by the hypothesis that $u = 0$. Similarly, if each $g_i \equiv \Sigma b_{ij} x_j \bmod \mathfrak{p}^2$, then

$$F_{\mathcal{O}}(\mathfrak{p}_S/\mathfrak{p}_S^2) = \{\det(b_{ij}): i \in I,\ \#I = r,\ I \subseteq \{1, \dots, s\}\}.$$

By (ii) again we see that $\det(a_{ij}) = \det(b_{ij})$ as ideals of $\mathcal{O}$ for some choice I_0 of I. After renumbering we may assume that $I_0 = \{1, \dots, r\}$. Then each g_i $(i = 1, \dots, r)$ can be written $g_i = \Sigma r_{ij} f_i$ for some $r_{ij} \in \mathcal{O}[\![x_1, \dots, x_r]\!]$ and we have

$$\det(b_{ij}) \equiv \det(r_{ij}) \cdot \det(a_{ij}) \mod \mathfrak{p}.$$

Hence $\det(r_{ij})$ is a unit, whence (r_{ij}) is an invertible matrix. Thus the f_i's can be expressed in terms of the g_i's and so $S \simeq T$.

We can extend this to the case $u \neq 0$ by picking $x_1, \dots, x_{r-u}$ so that they generate $(\mathfrak{p}_T/\mathfrak{p}_T^2)^{\text{tors}}$. Then we can write each $f_i \equiv \sum_{i=1}^{r-u} a_{ij} x_j \bmod \mathfrak{p}^2$ and likewise for the g_i's. The argument is now just as before but applied to the Fitting ideals of $(\mathfrak{p}_T/\mathfrak{p}_T^2)^{\text{tors}}$. □

For the next proposition we continue to assume that $\mathcal{O}$ is a complete discrete valuation ring. Let T be a local $\mathcal{O}$-algebra which as a module is finite and free over $\mathcal{O}$. In addition, we assume the existence of an isomorphism of T-modules $T \xrightarrow{\sim} \operatorname{Hom}_{\mathcal{O}}(T, \mathcal{O})$. We call a local $\mathcal{O}$-algebra which is finite and free and satisfies this extra condition a Gorenstein $\mathcal{O}$-algebra (cf. §5 of [Ti1]). Now suppose that $\mathfrak{p}$ is a prime ideal of T such that $T/\mathfrak{p} \simeq \mathcal{O}$.

Let $\beta: T \to T/\mathfrak{p} \simeq \mathcal{O}$ be the natural map and define a principal ideal of T by

$$(\eta_T) = (\hat{\beta}(1))$$

where $\hat{\beta}: \mathcal{O} \to T$ is the adjoint of β with respect to perfect $\mathcal{O}$-pairings on $\mathcal{O}$ and T, and where the pairing of T with itself is T-bilinear. (By a perfect pairing on a free $\mathcal{O}$-module M of finite rank we mean a pairing $M \times M \to \mathcal{O}$ such that both the induced maps $M \to \operatorname{Hom}_{\mathcal{O}}(M, \mathcal{O})$ are isomorphisms. When $M = T$ we are thus requiring that this be an isomorphism of T-modules also.) The ideal (η_T) is independent of the pairings. Also T/η_T is torsion-free as an $\mathcal{O}$-module, as can be seen by applying $\operatorname{Hom}(\ ,\mathcal{O})$ to the sequence

$$0 \to \mathfrak{p} \to T \to \mathcal{O} \to 0,$$

to obtain a homomorphism $T/\eta_T \hookrightarrow \operatorname{Hom}(\mathfrak{p}, \mathcal{O})$. This also shows that $(\eta_T) = \operatorname{Ann} \mathfrak{p}$.

If we let $l(M)$ denote the length of an $\mathcal{O}$-module M, then

$$l(\mathfrak{p}/\mathfrak{p}^2) \geq l(\mathcal{O}/\overline{\eta_T})$$

(where we write $\overline{\eta_T}$ for $\beta(\eta_T)$) because $\mathfrak{p}$ is a faithful T/η_T-module. (For a brief account of the relevant properties of Fitting ideals see the appendix to [MW1].) Indeed, writing $F_R(M)$ for the Fitting ideal of M as an R-module, we have

$$F_{T/\eta_T}(\mathfrak{p}) = 0 \Rightarrow F_T(\mathfrak{p}) \subset (\eta_T) \Rightarrow F_{T/\mathfrak{p}}(\mathfrak{p}/\mathfrak{p}^2) \subset (\overline{\eta_T})$$

and we then use the fact that the length of an $\mathcal{O}$-module M is equal to the length of $\mathcal{O}/F_{\mathcal{O}}(M)$ as $\mathcal{O}$ is a discrete valuation ring. In particular when $\mathfrak{p}/\mathfrak{p}^2$ is a torsion $\mathcal{O}$-module then $\overline{\eta}_T \neq 0$.

We need a criterion for a Gorenstein $\mathcal{O}$-algebra to be a complete intersection. We will say that a local $\mathcal{O}$-algebra S which is finite and free over $\mathcal{O}$ is a complete intersection over $\mathcal{O}$ if there is an $\mathcal{O}$-algebra isomorphism $S \simeq \mathcal{O}[\![x_1, \ldots, x_r]\!]/(f_1, \ldots, f_r)$ for some r. Such a ring is necessarily a Gorenstein $\mathcal{O}$-algebra and $\{f_1, \ldots, f_r\}$ is necessarily a regular sequence. That (i) $\Rightarrow$ (ii) in the following proposition is due to Tate (see A.3, conclusion 4, in the appendix in [M Ro].)

Proposition 2. *Assume that $\mathcal{O}$ is a complete discrete valuation ring and that T is a local Gorenstein $\mathcal{O}$-algebra which is finite and free over $\mathcal{O}$ and*

that $\mathfrak{p}_T$ is a prime ideal of T such that $T/\mathfrak{p}_T \simeq \mathcal{O}$ and $\mathfrak{p}_T/\mathfrak{p}_T^2$ is a torsion $\mathcal{O}$-module. Then the following two conditions are equivalent:

(i) *T is a complete intersection over $\mathcal{O}$.*

(ii) *$l(\mathfrak{p}_T/\mathfrak{p}_T^2) = l(\mathcal{O}/\overline{\eta_T})$ as $\mathcal{O}$-modules.*

Proof. To prove that (ii) $\Rightarrow$ (i), pick a complete intersection S over $\mathcal{O}$ (so assumed finite and flat over $\mathcal{O}$) such that $\alpha\colon S \twoheadrightarrow T$ and such that $\mathfrak{p}_S/\mathfrak{p}_S^2 \simeq \mathfrak{p}_T/\mathfrak{p}_T^2$ where $\mathfrak{p}_S = \alpha^{-1}(\mathfrak{p}_T)$. The existence of such an S seems to be well known (cf. [Ti2, §6]) but here is an argument suggested by N. Katz and H. Lenstra (independently).

Write $T = \mathcal{O}[x_1, \dots, x_r]/(f_1, \dots, f_s)$ with $\mathfrak{p}_T$ the image in T of $\mathfrak{p} = (x_1, \dots, x_r)$. Since T is local and finite and free over $\mathcal{O}$, it follows that also $T \simeq \mathcal{O}[\![x_1, \dots, x_r]\!]/(f_1, \dots, f_s)$. We can pick $g_1, \dots, g_r$ such that $g_i = \Sigma a_{ij} f_j$ with $a_{ij} \in \mathcal{O}$ and such that

$$(f_1, \dots, f_s, \mathfrak{p}^2) = (g_1, \dots, g_r, \mathfrak{p}^2).$$

We then modify $g_1, \dots, g_r$ by the addition of elements $\{\alpha_i\}$ of $(f_1, \dots, f_s)^2$ and set $(g_1' = g_1 + \alpha_1, \dots, g_r' = g_r + \alpha_r)$. Since T is finite over $\mathcal{O}$, there exists an N such that for each i, x_i^N can be written in T as a polynomial $h_i(x_1, \dots, x_r)$ of total degree less than N. We can assume also that N is chosen greater than the total degree of g_i for each i. Set $\alpha_i = (x_i^N - h_i(x_1, \dots, x_r))^2$. Then set $S = \mathcal{O}[\![x_1, \dots, x_r]\!]/(g_1' \dots, g_r')$. Then S is finite over $\mathcal{O}$ by construction and also $\dim(S) \le 1$ since $\dim(S/\lambda) = 0$ where (λ) is the maximal ideal of $\mathcal{O}$. It follows that $\{g_1', \dots, g_r'\}$ is a regular sequence and hence that $\operatorname{depth}(S) = \dim(S) = 1$. In particular the maximal $\mathcal{O}$-torsion submodule of S is zero since it is also a finite length S-submodule of S.

Now $\mathcal{O}/(\bar\eta_S) \simeq \mathcal{O}/(\bar\eta_T)$, since $l(\mathcal{O}/(\bar\eta_S)) = l(\mathfrak{p}_S/\mathfrak{p}_S^2)$ by (i) $\Rightarrow$ (ii) and $l(\mathcal{O}/(\bar\eta_T)) = l(\mathfrak{p}_T/\mathfrak{p}_T^2)$ by hypothesis. Pick isomorphisms

$$T \simeq \operatorname{Hom}_{\mathcal{O}}(T, \mathcal{O}), \quad S \simeq \operatorname{Hom}_{\mathcal{O}}(S, \mathcal{O})$$

as T-modules and S-modules, respectively. The existence of the latter for complete intersections over $\mathcal{O}$ is well known; cf. conclusion 1 of Theorem A.3 of [M Ro]. Then we have a sequence of maps, in which $\hat\alpha$ and $\hat\beta$ denote the adjoints with respect to these isomorphisms:

$$\mathcal{O} \xrightarrow{\hat\beta} T \xrightarrow{\hat\alpha} S \xrightarrow{\alpha} T \xrightarrow{\beta} \mathcal{O}.$$

One checks that $\hat\alpha$ is a map of S-modules (T being given an S-action via α) and in particular that $\alpha \circ \hat\alpha$ is multiplication by an element t of T. Now $(\beta \circ \hat\beta) = (\bar\eta_T)$ in $\mathcal{O}$ and $(\beta \circ \alpha) \circ \widehat{(\beta \circ \alpha)} = (\bar\eta_S)$ in $\mathcal{O}$. As $(\bar\eta_S) = (\bar\eta_T)$ in $\mathcal{O}$, we have that t is a unit mod $\mathfrak{p}_T$ and hence that $\alpha \circ \hat\alpha$ is an isomorphism. It follows

that $S \simeq T$, as otherwise $S \simeq \ker \alpha \oplus \operatorname{im} \hat{\alpha}$ is a nontrivial decomposition as S-modules, which contradicts S being local. □

Remark. Lenstra has made an important improvement to this proposition by showing that replacing $\bar{\eta}_T$ by $\beta(\operatorname{ann} \mathfrak{p})$ gives a criterion valid for all local $\mathcal{O}$-algebras which are finite and free over $\mathcal{O}$, thus without the Gorenstein hypothesis.

PRINCETON UNIVERSITY, PRINCETON, NJ

REFERENCES

[AK] A. ALTMAN and S. KLEIMAN, *An Introduction to Grothendieck Duality Theory*, vol. 146, Springer Lecture Notes in Mathematics, 1970.

[BiKu] B. BIRCH and W. KUYK (eds.), *Modular Functions of One Variable* IV, vol. 476, Springer Lecture Notes in Mathematics, 1975.

[Bo] Boston, N., *Families of Galois representations—Increasing the ramification*, Duke Math. J. **66**, 357–367.

[BH] W. BRUNS and J. HERZOG, *Cohen-Macaulay Rings*, Cambridge University Press, 1993.

[BK] S. BLOCH and K. KATO, *L-Functions and Tamagawa Numbers of Motives*, The Grothendieck Festschrift, Vol. 1 (P. Cartier et al. eds.), Birkhäuser, 1990.

[BLR] N. BOSTON, H. LENSTRA, and K. RIBET, Quotients of group rings arising from two-dimensional representations, C. R. Acad. Sci. Paris **t312**, Ser. 1 (1991), 323–328.

[CF] J. W. S. CASSELS and A. FRÖLICH (eds.), *Algebraic Number Theory*, Academic Press, 1967.

[Ca1] H. CARAYOL, Sur les représentations p-adiques associées aux formes modulaires de Hilbert, Ann. Sci. Ec. Norm. Sup. **IV**, Ser. 19 (1986), 409–468.

[Ca2] ———, Sur les représentations galoisiennes modulo ℓ attachées aux formes modulaires, Duke Math. J. **59** (1989), 785–801.

[Ca3] ———, Formes modulaires et représentations Galoisiennes à valeurs dans un anneau local complet, in *p-Adic Monodromy and the Birch-Swinnerton-Dyer Conjecture* (eds. B. Mazur and G. Stevens), Contemp. Math., vol. 165, 1994.

[CPS] E. CLINE, B. PARSHALL, and L. SCOTT, Cohomology of finite groups of Lie type I, Publ. Math. IHES **45** (1975), 169–191.

[CS] J. COATES and C. G. SCHMIDT, Iwasawa theory for the symmetric square of an elliptic curve, J. reine und angew. Math. **375/376** (1987), 104–156.

[CW] J. COATES and A. WILES, On p-adic L-functions and elliptic units, Ser. A26, J. Aust. Math. Soc. (1978), 1–25.

[Co] R. COLEMAN, Division values in local fields, Invent. Math. **53** (1979), 91–116.

[DR] P. DELIGNE and M. RAPOPORT, *Schémas de modules de courbes elliptiques*, in Springer Lecture Notes in Mathematics, Vol. 349, 1973.

[DS] P. DELIGNE and J-P. SERRE, Formes modulaires de poids 1, Ann. Sci. Ec. Norm. Sup. **IV**, Ser. 7 (1974), 507–530.

[Dia] F. DIAMOND, The refined conjecture of Serre, to appear in Proc. 1993 Hong Kong Conf. on Modular Forms and Elliptic Curves.

[Di] L. E. DICKSON, *Linear Groups with an Exposition of the Galois Field Theory*, Teubner, Leipzig, 1901.

[Dr] V. DRINFELD, Two-dimensional ℓ-adic representations of the fundamental group of a curve over a finite field and automorphic forms on GL(2), Am. J. Math. **105** (1983), 85–114.

[E1] B. EDIXHOVEN, The weight in Serre's conjecture on modular forms, Invent. Math. **109** (1992), 563–594.

[E2] ———, L'action de l'algèbre de Hecke sur les groupes de composantes des jacobiennes des courbes modulaires est "Eisenstein", in *Courbes Modulaires et Courbes de Shimura*, Astérisque **196-197** (1991), 159–170.

[Fl] M. FLACH, A finiteness theorem for the symmetric square of an elliptic curve, Invent. Math. **109** (1992), 307–327.

[Fo] J.-M. FONTAINE, Sur certains types de représentations p-adiques du groupe de Galois d'un corps local; construction d'un anneau de Barsotti-Tate, Ann. of Math. **115** (1982), 529–577.

[Fr] G. FREY, Links between stable elliptic curves and certain diophantine equations, Annales Universitatis Saraviensis **1** (1986), 1–40.

[Gre1] R. GREENBERG, Iwasawa theory for p-adic representations, Adv. St. Pure Math. **17** (1989), 97–137.

[Gre2] ———, On the structure of certain Galois groups, Invent. Math. **47** (1978), 85–99.

[Gro] B. H. GROSS, A tameness criterion for Galois representations associated to modular forms mod p, Duke Math. J. **61** (1990), 445–517.

[Guo] L. GUO, General Selmer groups and critical values of Hecke L-functions, Math. Ann. **297** (1993), 221–233.

[He] Y. HELLEGOUARCH, Points d'ordre $2p^h$ sur les courbes elliptiques, Acta Arith. **XXVI** (1975), 253–263.

[Hi1] H. HIDA, Iwasawa modules attached to congruences of cusp forms, Ann. Sci. Ecole Norm. Sup. (4) **19** (1986), 231–273.

[Hi2] ———, Theory of p-adic Hecke algebras and Galois representations, Sugaku Expositions **2-3** (1989), 75–102.

[Hi3] ———, Congruences of cusp forms and special values of their zeta functions, Invent. Math. **63** (1981), 225–261.

[Hi4] ———, On p-adic Hecke algebras for GL_2 over totally real fields, Ann. of Math. **128** (1988), 295–384.

[Hu] B. HUPPERT, *Endliche Gruppen I*, Springer-Verlag, 1967.

[Ih] Y. IHARA, On modular curves over finite fields, in Proc. Intern. Coll. on discrete subgroups of Lie groups and application to moduli, Bombay, 1973, pp. 161–202.

[Iw1] K. IWASAWA, *Local Class Field Theory*, Oxford University Press, Oxford, 1986.

[Iw2] ———, On $\mathbf{Z}_l$-extensions of algebraic number fields, Ann. of Math. **98** (1973), 246–326.

[Ka] N. KATZ, A result on modular forms in characteristic p, in *Modular Functions of One Variable* V, Springer L. N. M. **601** (1976), 53–61.

[Ku1] E. KUNZ, *Introduction to Commulative Algebra and Algebraic Geometry*, Birkhaüser, 1985.

[Ku2] ———, Almost complete intersections are not Gorenstein, J. Alg. **28** (1974), 111–115.

[KM] N. KATZ and B. MAZUR, *Arithmetic Moduli of Elliptic Curves*, Ann. of Math. Studies **108**, Princeton University Press, 1985.

[La] R. LANGLANDS, *Base Change for* GL (2), Ann. of Math. Studies, Princeton University Press **96**, 1980.

[Li] W. LI, Newforms and functional equations, Math. Ann. **212** (1975), 285–315.

[Liv] R. LIVNÉ, On the conductors of mod ℓ Galois representations coming from modular forms, J. of No. Th. **31** (1989), 133–141.

[Ma1] B. MAZUR, Deforming Galois representations, in *Galois Groups over* **Q**, vol. 16, MSRI Publications, Springer, New York, 1989.

[Ma2] ———, Modular curves and the Eisensten ideal, Publ. Math. IHES **47** (1977), 133–186.

[Ma3] ———, Rational isogenies of prime degree, Invent. Math. **44** (1978), 129–162.

[M Ri] B. MAZUR and K. RIBET, Two-dimensional representations in the arithmetic of modular curves, *Courbes Modulaires et Courbes de Shimura*, Astérisque **196-197** (1991), 215–255.

[M Ro] B. MAZUR and L. ROBERTS, Local Euler characteristics, Invent. Math. **9** (1970), 201–234.

[MT] B. MAZUR and J. TILOUINE, Représentations galoisiennes, differentielles de Kähler et conjectures principales, Publ. Math. IHES **71** (1990), 65–103.

[MW1] B. MAZUR and A. WILES, Class fields of abelian extensions of **Q**, Invent. Math. **76** (1984), 179–330.

[MW2] ———, On p-adic analytic families of Galois representations, Comp. Math. **59** (1986), 231–264.

[Mi1] J. S. MILNE, Jacobian varieties, in *Arithmetic Geometry* (Cornell and Silverman, eds.), Springer-Verlag, 1986.

[Mi2] ———, *Arithmetic Duality Theorems*, Academic Press, 1986.

[Ram] R. RAMAKRISHNA, On a variation of Mazur's deformation functor, Comp. Math. **87** (1993), 269–286.

[Ray1] M. RAYNAUD, Schémas en groupes de type $(p, p, \dots, p)$, Bull. Soc. Math. France **102** (1974), 241–280.

[Ray2] ———, Spécialisation du foncteur de Picard, Publ. Math. IHES **38** (1970), 27–76.

[Ri1] K. A. RIBET, On modular representations of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms, Invent. Math. **100** (1990), 431–476.

[Ri2] ———, Congruence relations between modular forms, Proc. Int. Cong. of Math. **17** (1983), 503–514.

[Ri3] ———, Report on mod l representations of $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, Proc. of Symp. in Pure Math. **55** (1994), 639–676.

[Ri4] ———, Multiplicities of p-finite mod p Galois representations in $J_0(Np)$, Boletin da Sociedade Brasileira de Matematica, Nova Serie **21** (1991), 177–188.

[Ru1] K. RUBIN, Tate-Shafarevich groups and L-functions of elliptic curves with complex multiplication, Invent. Math. **89** (1987), 527–560.

[Ru2] ———, The 'main conjectures' of Iwasawa theory for imaginary quadratic fields, Invent. Math. **103** (1991), 25–68.

[Ru3] ———, Elliptic curves with complex multiplication and the conjecture of Birch and Swinnerton-Dyer, Invent. Math. **64** (1981), 455–470.

[Ru4] ———, More 'main conjectures' for imaginary quadratic fields, CRM Proceedings and Lecture Notes, 4, 1994.

[Sch] M. SCHLESSINGER, Functors on Artin rings, Trans. A.M.S. **130** (1968), 208–222.

[Scho] R. SCHOOF, The structure of the minus class groups of abelian number fields, in Seminaire de Théorie des Nombres, Paris (1988–1989), Progress in Math. **91**, Birkhauser (1990), 185–204.

[Se] J-P. SERRE, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$, Duke Math. J. **54** (1987), 179–230.

[de Sh] E. DE SHALIT, *Iwasawa Theory of Elliptic Curves with Complex Multiplication*, Persp. in Math., Vol. 3, Academic Press, 1987.

[Sh1] G. SHIMURA, *Introduction to the Arithmetic Theory of Automorphic Functions*, Iwanami Shoten and Princeton University Press, 1971.

[Sh2] ———, On the holomorphy of certain Dirichlet series, Proc. London Math. Soc. (3) **31** (1975), 79–98.

[Sh3] ———, The special values of the zeta function associated with cusp forms, Comm. Pure and Appl. Math. **29** (1976), 783–804.

[Sh4] ———, On elliptic curves with complex multiplication as factors of the Jacobians of modular function fields, Nagoya Math. J. **43** (1971), 199–208.

[Ta] J. TATE, p-divisible groups, Proc. Conf. on Local Fields, Driebergen, 1966, Springer-Verlag, 1967, pp. 158–183.

[Ti1] J. TILOUINE, Un sous-groupe p-divisible de la jacobienne de $X_1(Np^r)$ comme module sur l'algebre de Hecke, Bull. Math. Soc. France **115** (1987), 329–360.

[Ti2] ———, Théorie d'Iwasawa classique et de l'algèbre de Hecke ordinaire, Comp. Math. **65** (1988), 265–320.

[Tu] J. TUNNELL, Artin's conjecture for representations of octahedral type, Bull. A.M.S. **5** (1981), 173–175.

[TW] R. TAYLOR and A. WILES, Ring theoretic properties of certain Hecke algebras, next paper, this issue.

[We] A. WEIL, Uber die Bestimmung Dirichletscher Reihen durch Funktionalgleichungen, Math. Ann. **168** (1967), 149–156.

[Wi1] A. WILES, On ordinary λ-adic representations associated to modular forms, Invent. Math. **94** (1988), 529–573.

[Wi2] ———, On p-adic representations for totally real fields, Ann. of Math. **123** (1986), 407–456.

[Wi3] ———, Modular curves and the class group of $\mathbf{Q}(\zeta_p)$, Invent. Math. **58** (1980), 1–35.

[Wi4] ———, The Iwasawa conjecture for totally real fields, Ann. of Math. **131** (1990), 493–540.

[Win] J. P. WINTENBERGER, Structure galoisienne de limites projectives d'unitées locales, Comp. Math. **42** (1981), 89–103.

(Received October 14, 1994)